

EXHIBIT 2



US010337001B2

(12) **United States Patent**
Ryan et al.

(10) **Patent No.:** US 10,337,001 B2
(b4) **Date of Patent:** Jul. 2, 2019

(54) **GUIDE RNA WITH CHEMICAL MODIFICATIONS**(71) Applicant: **Agilent Technologies, Inc.**, Santa Clara, CA (US)(72) Inventors: **Daniel E. Ryan**, San Francisco, CA (US); **Douglas J. Dellinger**, Boulder, CO (US); **Jeffrey R. Sampson**, San Jose, CA (US); **Robert Kaiser**, Santa Clara, CA (US); **Joel Myerson**, Berkeley, CA (US)(73) Assignee: **AGILENT TECHNOLOGIES, INC.**, Santa Clara, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **15/607,295**(22) Filed: **May 26, 2017**(65) **Prior Publication Data**

US 2018/0051281 A1 Feb. 22, 2018

Related U.S. Application Data

(63) Continuation of application No. 14/757,204, filed on Dec. 3, 2015.

(60) Provisional application No. 62/256,095, filed on Nov. 16, 2015, provisional application No. 62/146,189, filed on Apr. 10, 2015, provisional application No. 62/087,211, filed on Dec. 3, 2014.

(51) **Int. Cl.**

C12N 15/113	(2010.01)
C12N 15/11	(2006.01)
C12N 9/22	(2006.01)
C07H 21/02	(2006.01)
C12N 15/90	(2006.01)
C12Q 1/6876	(2018.01)

(52) **U.S. Cl.**

CPC	C12N 15/111 (2013.01); C07H 21/02 (2013.01); C12N 9/22 (2013.01); C12N 15/113 (2013.01); C12N 15/907 (2013.01); C12Q 1/6876 (2013.01); C12N 2310/10 (2013.01); C12N 2310/20 (2017.05); C12N 2310/312 (2013.01); C12N 2310/315 (2013.01); C12N 2310/321 (2013.01); C12N 2310/322 (2013.01); C12N 2310/323 (2013.01); C12N 2310/3231 (2013.01); C12N 2310/333 (2013.01); C12N 2310/335 (2013.01); C12N 2310/346 (2013.01); C12N 2310/3517 (2013.01); C12N 2310/531 (2013.01); C12N 2320/51 (2013.01); C12N 2320/53 (2013.01); C12N 2330/31 (2013.01)
-----------	--

(58) **Field of Classification Search**

None

See application file for complete search history.

(56)

References Cited

U.S. PATENT DOCUMENTS

5,032,401 A	7/1991	Jamas et al.
5,607,677 A	3/1997	Jamas et al.
7,371,580 B2	5/2008	Yakhini et al.
8,202,983 B2	6/2012	Dellinger et al.
8,906,616 B2 *	12/2014	Zhang C12N 15/63 435/440
2005/0281781 A1	12/2005	Ostroff
2010/0076183 A1	3/2010	Dellinger et al.
2014/0090113 A1	3/2014	Cogan et al.
2014/0090116 A1	3/2014	Ainley et al.
2014/0170753 A1	6/2014	Zhang
2014/0242664 A1 *	8/2014	Zhang C12N 15/902 435/188
2014/0273037 A1	9/2014	Wu
2014/0273226 A1	9/2014	Wu
2014/0273232 A1	9/2014	Zhang et al.
2014/0273233 A1	9/2014	Chen et al.
2014/0273235 A1	9/2014	Voytas et al.
2014/0294773 A1	10/2014	Brouns et al.
2014/0295555 A1	10/2014	Joung et al.
2014/0295557 A1	10/2014	Joung et al.
2014/0309487 A1	10/2014	Lee et al.
2014/0310828 A1	10/2014	Lee et al.
2014/0349400 A1	11/2014	Jakimo et al.
2014/0357523 A1	12/2014	Zeiner et al.
2014/0364333 A1	12/2014	Wu et al.
2015/0044191 A1	2/2015	Liu et al.
2015/0044192 A1	2/2015	Liu et al.
2015/0044772 A1	2/2015	Zhao
2015/0064149 A1	3/2015	West et al.
2015/0064708 A1	3/2015	Sastry-Dent et al.
2015/0067921 A1	3/2015	Cogan et al.
2015/0071906 A1	3/2015	Liu et al.
2015/0071946 A1	3/2015	Solyom et al.
2015/0079680 A1	3/2015	Bradley et al.
2015/0098954 A1	4/2015	Hyde et al.
2015/0128307 A1	5/2015	Sastry-Dent et al.
2015/0128308 A1	5/2015	Sastry-Dent et al.
2015/0128309 A1	5/2015	Sastry-Dent et al.
2015/0133315 A1	5/2015	Jacobson et al.
2015/0140664 A1	5/2015	Byrne et al.
2015/0156996 A1	6/2015	Fahrenkrug et al.
2015/0315252 A1	11/2015	Haugwitz et al.
2015/0315576 A1	11/2015	Caliendo et al.
2015/0322432 A1	11/2015	Anderson et al.
2015/0344836 A1	12/2015	Finer et al.

(Continued)

FOREIGN PATENT DOCUMENTS

EP	2800811 A1	11/2014
EP	2892321 A2	7/2015

(Continued)

OTHER PUBLICATIONS

Jinek et al., "RNA-programmed genome editing in human cells," eLife (2013), 2, e00471.

(Continued)

Primary Examiner — Channing S Mahatan

(57) **ABSTRACT**

The present invention relates to modified guide RNAs and their use in clustered, regularly interspaced, short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems.

30 Claims, 17 Drawing Sheets**Specification includes a Sequence Listing.**

US 10,337,001 B2

Page 2

(56)	References Cited					
U.S. PATENT DOCUMENTS						
2016/0030477 A1	2/2016	Conway et al.	WO	2015052231 A2	4/2015	
2016/0046959 A1	2/2016	Landel et al.	WO	2015053995 A1	4/2015	
2016/0090622 A1	3/2016	Liu et al.	WO	2015054253 A1	4/2015	
2016/0138027 A1	5/2016	Gan et al.	WO	2015054375 A2	4/2015	
2016/0168592 A1	6/2016	Church et al.	WO	2015066636 A2	5/2015	
2016/0206566 A1	7/2016	Lu et al.	WO	2015066637 A1	5/2015	
2016/0298097 A1	10/2016	Chavez et al.	WO	2015070083 A1	5/2015	
2017/0166893 A1	6/2017	Doudna et al.	WO	2015075056 A1	5/2015	
2017/0298383 A1	10/2017	Albertsen et al.	WO	2015089277 A1	6/2015	
FOREIGN PATENT DOCUMENTS						
EP	2893006 A1	7/2015	WO	2015089351 A1	6/2015	
EP	3241902 B1	2/2018	WO	2015089354 A1	6/2015	
WO	2013126794 A1	8/2013	WO	2015089406 A1	6/2015	
WO	2013141680 A1	9/2013	WO	2015089419 A2	6/2015	
WO	WO2013142578	9/2013	WO	2015089427 A1	6/2015	
WO	2013176772	11/2013	WO	2015089462 A1	6/2015	
WO	WO2013176772	11/2013	WO	2015089465 A1	6/2015	
WO	WO2013176844	11/2013	WO	2015089473 A1	6/2015	
WO	2014039684	3/2014	WO	2015089486 A2	6/2015	
WO	2014039702	3/2014	WO	2014204728 A8	7/2015	
WO	2014065596 A1	5/2014	WO	2015184259 A1	12/2015	
WO	2014071219 A1	5/2014	WO	2015184262 A1	12/2015	
WO	2014089290 A1	6/2014	WO	2015184268 A1	12/2015	
WO	2014089513 A1	6/2014	WO	2015200334 A1	12/2015	
WO	2014089533 A2	6/2014	WO	2015200378 A1	12/2015	
WO	2014093595 A1	6/2014	WO	2015200555 A2	12/2015	
WO	2014093622 A2	6/2014	WO	2015200725 A1	12/2015	
WO	2014093635 A1	6/2014	WO	2014204723 A9	2/2016	
WO	2014093655 A2	6/2014	WO	2016022363 A2	2/2016	
WO	2014093661 A2	6/2014	WO	2014204724 A9	3/2016	
WO	2014093688 A1	6/2014	WO	2016033246 A1	3/2016	
WO	2014093694 A1	6/2014	WO	2016033315 A2	3/2016	
WO	2014093701 A1	6/2014	WO	2016046288 A1	3/2016	
WO	2014093709 A1	6/2014	WO	2016049024 A2	3/2016	
WO	2014093712 A1	6/2014	WO	2016049163 A2	3/2016	
WO	2014093718 A1	6/2014	WO	2016049251 A1	3/2016	
WO	2014099744 A1	6/2014	WO	2016049657 A1	3/2016	
WO	2014099750 A2	6/2014	WO	2016057755 A1	4/2016	
WO	2014144288 A1	9/2014	WO	2016057821 A2	4/2016	
WO	2014145599 A2	9/2014	WO	2016057835 A2	4/2016	
WO	2014150624 A1	9/2014	WO	2016057951 A2	4/2016	
WO	2014152432 A2	9/2014	WO	2016069282 A1	5/2016	
WO	2014153118 A1	9/2014	WO	2016069283 A1	5/2016	
WO	WO2014144592	9/2014	WO	2016070037 A2	5/2016	
WO	WO2014144761	9/2014	WO	2016073079 A2	5/2016	
WO	2014039970 A9	10/2014	OTHER PUBLICATIONS			
WO	2014159719 A1	10/2014	Lujambo et al., "Genetic Unmasking of an Epigenetically Silenced microRNA in Human Cancer Cells," (Spanish National Cancer Centre) Cancer Res. Feb. (2007).			
WO	2014165349 A1	10/2014	Mali et al., "CRISPR transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering," Nat. Biotechnol. (2013), 31, 833-8.			
WO	2014165825 A2	10/2014	Larson et al., Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene, Science, vol. 332, No. 6028, Apr. 22, 2011, 475-478.			
WO	2014172458 A1	10/2014	Bisaria et al., "Lessons from enzyme kinetics reveal specificity principles for RNA-guided nucleases in RNA interference and CRISPR-based genome editing," Cell Syst., 4:21-29, (2017).			
WO	2014172470 A2	10/2014	Belfort et al., "Homing Endonucleases: Keeping the House in Order", Nucleic Acids Res., vol. 25, pp. 3379-3388, (1997).			
WO	2014186585 A2	11/2014	Cho et al., "Targeted Genome Engineering in Human Cells with the cas9-mg-guided endonuclease", Nat. Biotechnol. Mar. 2013.			
WO	2014191128 A1	12/2014	Cho et al., "Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases," Genome Res., vol. 24, pp. 0132-0141, (2014).			
WO	2014191518 A1	12/2014	Dellinger et al., "Solid-Phase Chemical Synthesis of Phosphonoacetate and Thiophosphonoacetate Oligodeoxynucleotides", J. Am. Chem. Soc., vol. 125, pp. 940-950, (2003).			
WO	2014191521 A2	12/2014	Dellinger et al., "Streamlined Process for the Chemical Synthesis of RNA Using 2'-O-Thionocarbamate-Protected Nucleoside Phosphoramidites in the Solid Phase", J. Am. Chem. Soc., vol. 133, pp. 11540-11556, (2011).			
WO	2014197568 A2	12/2014	Doyon et al., "Enhancing Zinc-Finger-Nuclease Activity with Improved Obligate Heterodimeric Architectures", Nat. Methods, vol. 8, pp. 74-81, (2011).			
WO	2014201015 A2	12/2014				
WO	2014204725 A1	12/2014				
WO	2014204726 A1	12/2014				
WO	2015006294 A2	1/2015				
WO	2015006437 A1	1/2015				
WO	2015006498 A2	1/2015				
WO	2015013583 A2	1/2015				
WO	2015026883 A1	2/2015				
WO	2015026885 A1	2/2015				
WO	2015026886 A1	2/2015				
WO	2015030881 A1	3/2015				
WO	2015033293 A1	3/2015				
WO	2015035136 A2	3/2015				
WO	2015035139 A2	3/2015				
WO	2015035162 A2	3/2015				
WO	2015040075 A1	3/2015				
WO	2015052133 A1	4/2015				

US 10,337,001 B2

Page 3

(56)

References Cited

OTHER PUBLICATIONS

- El-Sagheer, A.H., et al. "Click chemistry with DNA", Chem. Soc. Rev., vol. 39, pp. 1388-1405, (2010).
- Frock et al., "Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases," Nat. Biotechnol., vol. 33, pp. 179-186, (2015).
- Gao et al., "Single Cas9 nickase induced generation of NRAMP1 knockin cattle with reduced off-target effects," Genome Biol., vol. 18, p. 13, (2017).
- Gasiunas et al., "Cas9—crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria", Proc. Natl. Acad. Sci. USA, 109:39, (2012).
- Fu et al., "High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells," Nat. Biotechnol., vol. 31, pp. 822-826, (2013).
- <http://crispr.mit.edu>.
- <http://www.genome.net/Cas-OFFinder>.
- <https://cm.jefferson.edu/Off-Spotter>.
- Piccirilli, J. A., et al., "Enzymatic Incorporation of a New Base pair into DNA and RNA Extends the Genetic Alphabet." Nature, vol. 343, p. 33, (1990).
- Schneider et al. "Information Content of Binding Sites on Nucleotide Sequences" J. Mol. Biol., 188, 415-431 (1986).
- Jinek et al., "A Programmable Dual-RNA—Guided DNA Endonuclease in Adaptive Bacterial Immunity", Science, vol. 337, pp. 816-821, (2012).
- Krueger et al., "Synthesis and Properties of Size-Expanded DNAs: Toward Designed, Functional Genetic Systems", Acc. Chem. Res., vol. 40, pp. 141-150, (2007).
- Lahoud et al., "Enzymatic Synthesis of Structure-Free DNA with Pseudo-Complementary Properties", Nucl. Acids Res. vol. 36, 3409-19 (2008).
- Lange et al., "Classical Nuclear Localization Signals: Definition, Function, and Interaction with Importin α^* ", J. Biol. Chem., vol. 282, pp. 5101-5105, (2007).
- Kuznetsova, S.A., "Synthesis and properties of DNA duplexes containing hydrocarbon bridges instead of a nucleoside residue", Bioorganicheskaja khimiia/Akademija nauk SSSR, vol. 14:12, pp. 1656-1662, (1988).
- Lujambio et al., "Genetic Unmasking of an Epigenetically Silenced microRNA in Human Cancer Cells", Cancer Res., (2007).
- Miller et al., "An improved zinc-finger nuclease architecture for highly specific genome editing", Nat. Biotechnol. vol. 25, pp. 778-785, (2007).
- Mojibul, H.M., et al., "DNA-associated click chemistry", Sci. China Chem., vol. 57:2, pp. 215-231, 2014.
- New England Biolabs Catalog.
- Kumar et al., "Template-Directed Oligonucleotide Strand Ligation, Covalent Intramolecular DNA Circularization and Catenation Using Click Chemistry", J. Am. Chem. Soc., vol. 129, pp. 6859-6864, (2007).
- Ran et al., "In vivo genome editing using *Staphylococcus aureus* Cas9", Nature, vol. 520, pp. 186-191, (2015).
- Rappaport, H. P., "Replication of the Base Pair 6-Thioguanine/S-Methyl-2-pyrimidinone with the Large Klenow Fragment of *Escherichia coli* DNA Polymerase I", Biochemistry, vol. 32, p. 3047, (1993).
- Shabarova, "Synthesis and properties of DNA duplexes containing hydrocarbon bridges instead of a nucleoside residue", Bioorg. Khim., vol. 14:12, pp. 1656-1662, (1988).
- Gori et al., "Delivery and specificity of CRISPR/Cas9 genome editing technologies for human gene therapy," Hum. Gene Ther., vol. 26, pp. 443-451, (2015).
- Szczepk et al., "Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases", Nat. Biotechnol., vol. 25, pp. 786-793, (2007).
- Threlfall et al., "Synthesis and biological activity of phosphonoacetate- and thiophosphonoacetate-modified 2'-O-methyl oligoribonucleotides", Org. Biomol. Chem., vol. 10, pp. 746-754, (2012).
- Tietze et al., "Squaric Acid Diethyl Ester: A New Coupling Reagent for the formation of Drug Biopolymer Conjugates. Synthesis of Squaric Acid Ester Amides and Diamides", Chem. Ber., vol. 124, pp. 1215-1221, (1991).
- Fu et al., "Distinct patterns of Cas9 mismatch tolerance in vitro and in vivo," Nucleic Acids Res., vol. 44, pp. 5365-5377, (2016).
- Yamada, et al., "Synthesis and Biochemical Evaluation of Phosphonoformate Oligodeoxyribonucleotides", J. Am. Chem. Soc., vol. 128:15, pp. 5251-5261, (2006).
- Fu et al., "Improving CRISPR-Cas nuclease specificity using truncated guide RNAs," Nat. Biotechnol., vol. 32, pp. 279-284, (2014).
- Cradick et al., "CRISPR/Cas9 systems targeting β -globin and CCR5 genes have substantial off-target activity," Nucleic Acids Res., vol. 41, pp. 9584-9592, (2013).
- Davis et al., "Small molecule-triggered Cas9 protein with improved genome-editing specificity," Nat. Chem. Biol., vol. 11, pp. 316-318, (2013).
- Doench et al., "Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9," Nat. Biotechnol., vol. 34, pp. 184-191 (2016).
- Karvelis, et al., "crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*", RNA Biology, vol. 10, No. 5, May 2013,841-851.
- PCT, "International Search Report and Written Opinion dated Apr. 19, 2016", Application No. PCT/US2015/000143.
- Nishimasu et al., "Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA," Cell (2014) 156,1-15.
- Yang et al., "Artificially expanded genetic information system: a new base pair with an alternative hydrogen bonding pattern," Nucleic Acids Res. (2006), 34, 6095-101.
- Briner et al., "Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality," Mol. Cell (2014),56,333-9.
- Schechner et al., "CRISPR Display: A modular method for locus-specific targeting of long noncoding RNAs and synthetic RNA devices in vivo," Nat. Methods (2015), 12(7): 664-670.
- Chen et al., "Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System," Cell (2013),155,1479-1491.
- Mali et al., "RNA-Guided Human Genome Engineering via Cas9," Science (2013) , 339:6121, 823-6.
- Cong et al., "Multiplex Genome Engineering Using CRISPR/Cas Systems," Science (2013) , 339:6121, 819-23.
- Hwang et al., "Efficient In Vivo Genome Editing Using RNA-Guided Nucleases," Nat. Biotechnol. (2013) 31 :3, 227-9.
- Guilinger et al., "Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification," Nat. Biotechnol. 32, 577-582 (2014).
- Haeussler et al., "Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR," Genome Biol. 17, 148 (2016).
- Havlicek et al., "Re-engineered RNA-guided FokI-nucleases for improved genome editing in human cells," Mol. Ther. 25, 342-355 (2017).
- Hendel et al., "Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing," Cell Rep. 7, 293-305 (2014).
- Hendel et al., "Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells," Nat. Biotechnol. 33, 985-989 (2015).
- Hruscha et al., "Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish," Development 140, 4982-4987 (2013).
- Hsu et al., "DNA targeting specificity of RNA-guided Cas9 nucleases," Nat. Biotechnol. 31, 827-832 (2013).
- Hu et al., "Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing," Nat. Protoc. 11, 853-871 (2016).
- Iyer et al., "Off-target mutations are rare in Cas9-modified mice," Nat. Methods 12, 479 (2015).
- Jiang et al., "A Cas9-guide RNA complex preorganized for target DNA recognition," Science 348, 1477-1481 (2015).
- Kim et al., "Highly Efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins," Genome Res. 24, 1012-1019 (2014).

US 10,337,001 B2

Page 4

(56)

References Cited

OTHER PUBLICATIONS

- Kim et al., "Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells," *Nat. Methods* 12, 237-243 (2015).
- Kim et al., "Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq," *Genome Res.* 26, 406-415 (2016).
- Kim et al., "Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells," *Nat. Biotechnol.* 34, 863-868 (2016).
- Kleinstiver et al., "High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects," *Nature* 529, 490-495 (2016).
- Kleinstiver et al., "Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells," *Nat. Biotechnol.* 34, 869-874 (2016).
- Knight et al., "Dynamics of CRISPR-Cas9 genome interrogation in living cells," *Science* 350, 823-826 (2015).
- Komor et al., "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage," *Nature* 533, 420-424 (2016).
- Koo et al., "Measuring and reducing off-target activities of programmable nucleases including CRISPR-Cas9," *Mol. Cells* 38, 475-481 (2015).
- Kuscu et al., "Genome-wide analysis reveals characteristics of off-target sites bound by Cas9 endonuclease," *Nat. Biotechnol.* 32, 677-683 (2014).
- Lee et al., "Nuclease target site selection for maximizing on-target activity and minimizing off-target effects in genome editing," *Mol. Ther.* 24, 475-487 (2016).
- Ledford, "CRISPR, the Disruptor," *Nature* 522, 20-24 (2015).
- Lim et al., "Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease," *Nat. Commun.* 7, 13350 (2016).
- Lin et al., "CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences," *Nucleic Acids Res.* 42, 7473-7485 (2014).
- Ma et al., "CRISPR-Cas9 nuclear dynamics and target recognition in living cells," *J. Cell Biol.* 214, 529-537 (2016).
- Pattanayak et al., "High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity," *Nat. Biotechnol.* 31, 839-843 (2013).
- Peterson et al., "Genome-wide assessment of efficiency and specificity in CRISPR/Cas9 mediated multiple site targeting in *Arabidopsis*," *PLoS One* 11, e0162169 (2016).
- Polstein et al., "Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators," *Genome Res.* 25, 1158-1169 (2015).
- Qiu et al., "Mutation detection using Surveyor™ nuclease," *Biotechniques* 36, 702-707 (2004).
- Ran et al., "Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity," *Cell* 154, 1380-1389 (2013).
- Ren et al., "Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in *Drosophila*," *Cell Rep.* 9, 1151-1162 (2014).
- Semenova et al., "Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence," *Proc. Natl. Acad. Sci. USA* 108, 10098-10103 (2011).
- Shen et al., "Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects," *Nat. Methods* 11, 399-402 (2014).
- Singh et al., "Cas9-chromatin binding information enables more accurate CRISPR off-target prediction," *Nucleic Acids Res.* 43, e118 (2015).
- Singh et al., "Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9," *Nat. Commun.* 7, 12778 (2016).
- Slaymaker et al., "Rationally engineered Cas9 nucleases with improved specificity," *Science* 351, 84-88 (2016).
- Smith et al., "Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs," *Cell Stem Cell.* 15, 12-13 (2014).
- Tsai et al., "Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing," *Nat. Biotechnol.* 32, 569-576 (2014).
- Tsai et al., "Guide-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases," *Nat. Biotechnol.* 33, 187-197 (2015).
- Tsai & Joung, "Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases," *Nat. Rev. Genet.* 17, 300-312 (2016).
- Tycko et al., "Methods for optimizing CRISPR-Cas9 genome editing specificity," *Mol. Cell* 63, 355-370 (2016).
- Wang et al., "Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors," *Nat. Biotechnol.* 33, 175-178 (2015).
- Wu et al., "Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells," *Nat. Biotechnol.* 32, 670-676 (2014).
- Wu et al., "Target specificity of the CRISPR-Cas9 system," *Quant. Biol.* 2, 59-70 (2014).
- Wyvekens et al., "Dimeric CRISPR RNA-guided FokI-dCas9 nucleases directed by truncated gRNAs for highly specific genome editing," *Hum. Gene Ther.* 26, 425-431 (2015).
- Yang et al., "Targeted and genome-wide sequencing reveal single nucleotide variations impacting specificity of Cas9 in human stem cells," *Nat. Commun.* 5, 5507 (2014).
- Zetsche et al., "Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system," *Cell* 163, 759-771 (2015).
- Zischewski et al., "Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases," *Biotech. Adv.* 35, 95-104 (2017).
- Zheng et al., "Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence," *Sci. Rep.* 7, 40638 (2017).
- Zhang et al., "Evolution of Functional Six-Nucleotide DNA," *J. Am. Chem. Soc.*, 2015, 137 (21), 6734-6737.
- Deng, Wulan et al., "CASFISH: CRISPR/Cas9-Mediated In Situ Labeling of Genomic Loci in Fixed Cells," *Proceedings of the National Academy of Sciences*, vol. 112, No. 38, Sep. 22, 2015, 11870-11875.
- Extended European Search Report dated Jun. 13, 2018, Application No. 15866342.7, 14 pages.
- Rahdar, Meghdad et al., "Synthetic CRISPR RNA-Cas9-Guided Genome Editing Human Cells," *Proceedings of the National Academy of Sciences*, Nov. 16, 2015, 8 pages.

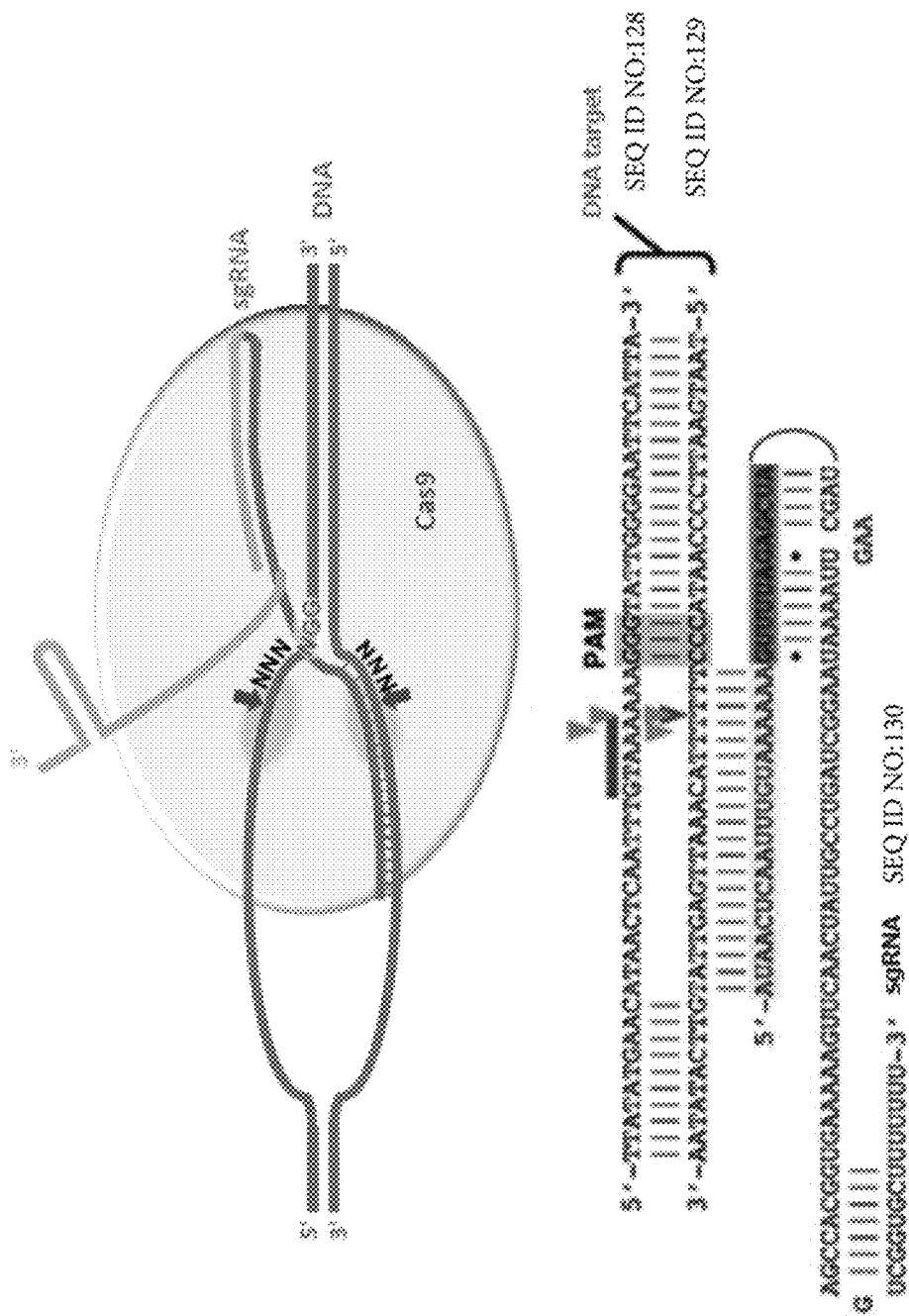
* cited by examiner

U.S. Patent

Jul. 2, 2019

Sheet 1 of 17

US 10,337,001 B2



二

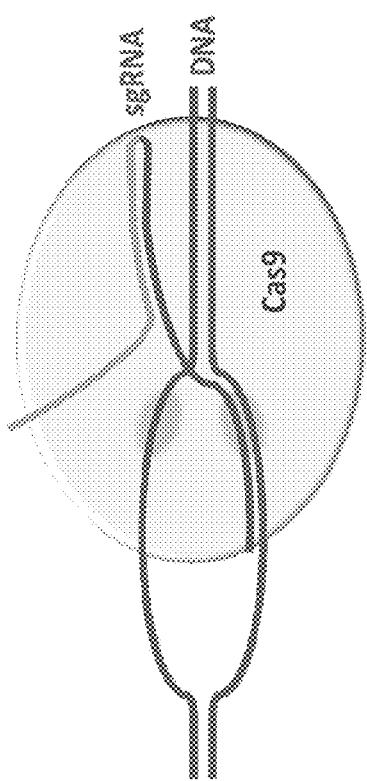
U.S. Patent

Jul. 2, 2019

Sheet 2 of 17

US 10,337,001 B2

Final concn in pmol/μl reaction	
Reaction volume	20 μl
Cas9 protein	40 nM (125 ng/20 μL rxn)
Guide RNA	50 nM
Target DNA	
TrisHCl pH 7.6	100 mM
NaCl	50 mM
MgCl ₂	10 mM



• employed version 1.0 sgRNA design

FIG. 2B

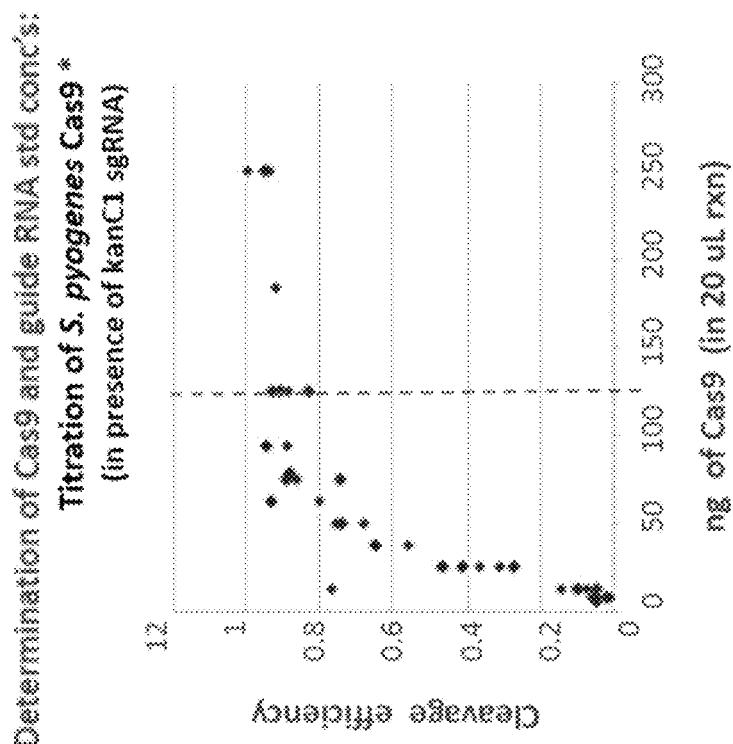
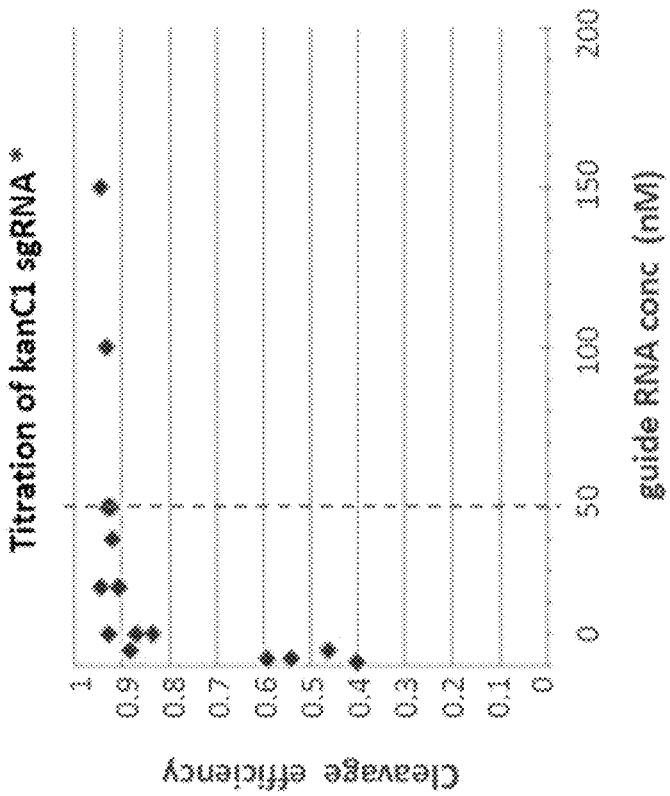
FIG. 2A

U.S. Patent

Jul. 2, 2019

Sheet 3 of 17

US 10,337,001 B2

**FIG. 2C****FIG. 2D**

U.S. Patent

Jul. 2, 2019

Sheet 4 of 17

US 10,337,001 B2

Standard cleavage conditions:

stock	volume (μ L)	conc.	stoich.
10X buffer	2.0	1x	
125 ng/ μ L Cas9 wt protein	1.0	40 nM	
1 μ M guide RNA	1.0	50 nM	
25 ng/ μ L linearized CTFA target	4.0	2.5 nM	
ddH ₂ O (DEPC-treated)	12.0		
			20
			1
			16
	20.0		

If every ng of Cas9 in our protein prep were active

In a pre-warmed SureCycler:

- (i) incubate at 37 °C for 30 min
- (ii) + RNase cocktail, incubate at 37 °C for 5 min, then at 70 °C for 15 min
- (iii) + Proteinase K, incubate at 37 °C for 15 min

Analyze crude products on Bioanalyzer.

FIG. 3

U.S. Patent

Jul. 2, 2019

Sheet 5 of 17

US 10,337,001 B2

三

4
6
L

U.S. Patent

Jul. 2, 2019

Sheet 6 of 17

US 10,337,001 B2

FIG. 4 (cont.)

U.S. Patent

Jul. 2, 2019

Sheet 7 of 17

US 10,337,001 B2

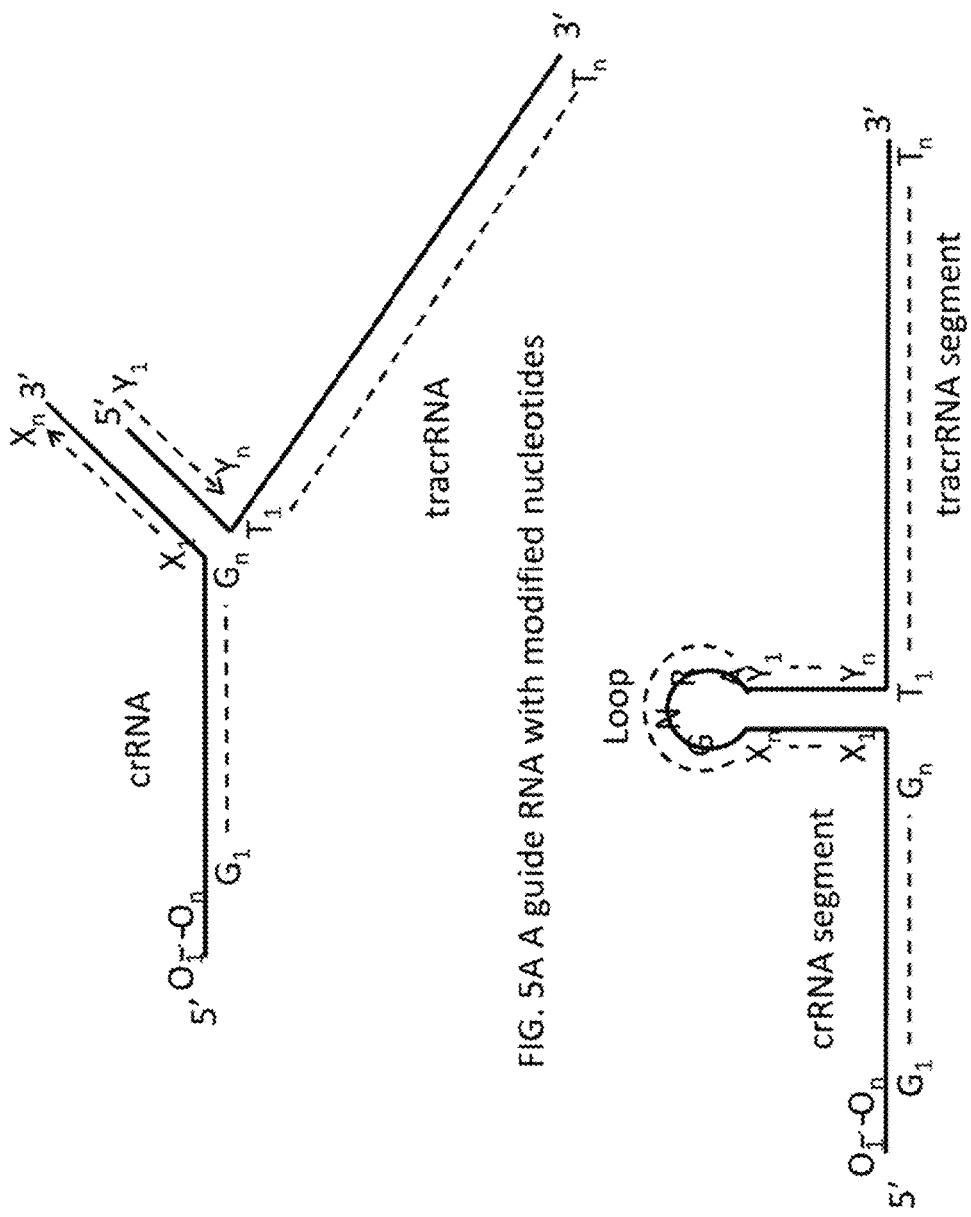


FIG. 5A A guide RNA with modified nucleotides

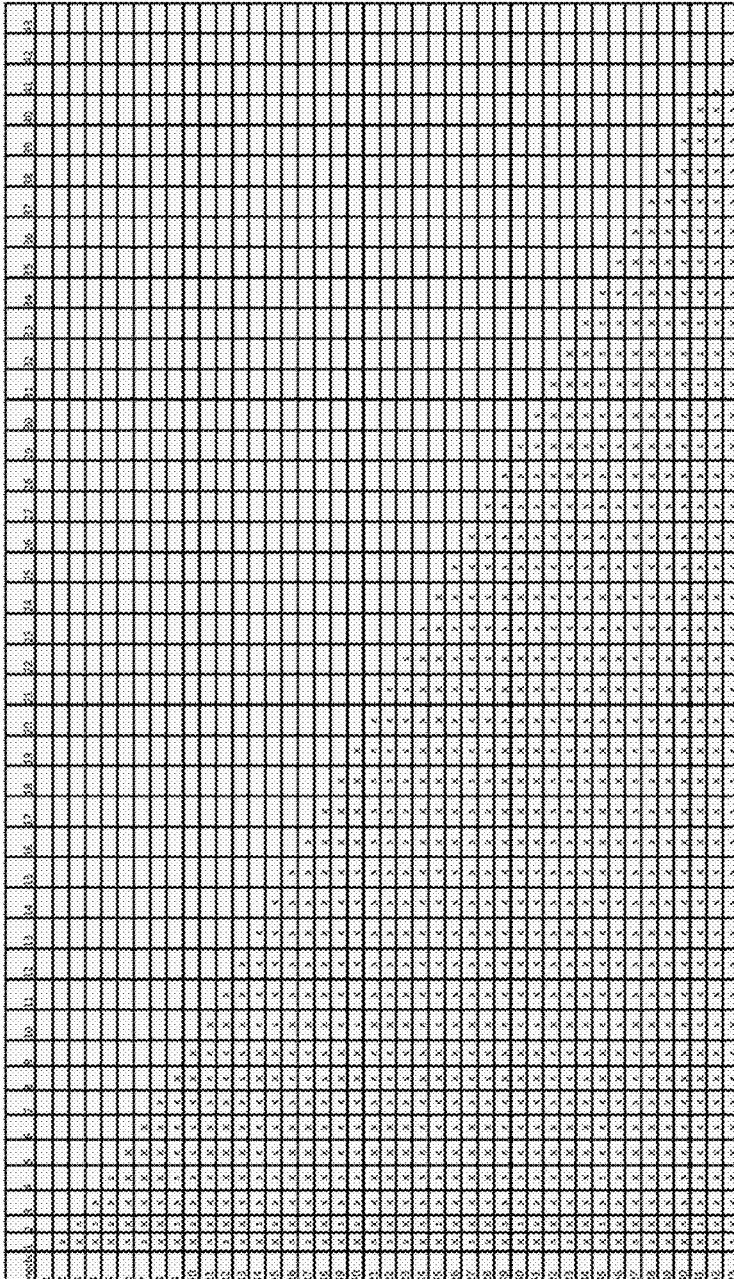
FIG. 5B A single guide RNA with modified nucleotides

U.S. Patent

Jul. 2, 2019

Sheet 8 of 17

US 10,337,001 B2



- | | | | | | | |
|------|--------------------------|------|--------------------|----|---------------------|---------------------------|
| I-1 | 2'-OMe | II-1 | O-PS linkage | 21 | 2-aminoepine | 39. Abasic nucleotide |
| I-2 | 2'-H | II-2 | P-ACE linkage | 22 | pseudouracil | 40. PEG |
| I-3 | 2'-ACN analog | II-3 | diP-ACE linkage | 23 | Hyperosanthine | 41. Hydrocarbon spacer |
| I-4 | 2'-Fluoro analog | II-4 | Methylphosphonate | 24 | 7-deazaG(A/C) | 42. Linker-dye |
| I-5 | 2'-antio analog | II-5 | Boranophosphonate | 25 | 7-deaza-8-azatA/G | 43. Spermine-modification |
| I-6 | 2'-azabutane analog | II-6 | Phosphorodithioate | 26 | 5-methyl U/C | |
| I-7 | 2'-Z-carbamino analog | II-7 | Z-thioU/C | 27 | 5-hydroxymethyl U/C | |
| I-8 | Z-U/A analog | II-8 | 4-thiouracil | 28 | 5,6-disubstituted | |
| I-9 | 2'-U/A analog | II-9 | 6-thioguanine | 29 | 5-propargyl U/C | |
| I-10 | 4'-chlorotriethyl methyl | | 2'-aminoadenine | | | 38. IsocC/G |

६

U.S. Patent

Jul. 2, 2019

Sheet 9 of 17

US 10,337,001 B2

Single mod	Sugar	Phosphorus linkage	Base modification*	Other
Double mod				
Sugar/Sugar	X	X	X	X
Sugar/p link	X	X	X	X
Sugar/Base	X	X	X	X
Sugar/other	X	X	X	X
p link/ p link	X	X	X	X
p link/ base	X	X	X	X
p link/other	X	X	X	X
Base/ Base	X	X	X	X
Base/other	X	X	X	X
other/other	X	X	X	X

*Base modifications includes Base Pair Modifications

Sugar modifications ("Sugar"): 2'-O-Methyl (=2'-OMe) {2'-OC₁-C₄ alkyl}, 2'-H, 2'-MOE {2'-OC₁-C₂ alkyl-O-C₁-C₃ alkyl}, 2'-F, 2'-amino, 2'-arabino, 2'-F-arabinosyl, 2'-LNA, 2'-UNNA, 4'-thioribosyl nucleotide.

Internucleotide linkage and 3' and/or 5' terminal nucleotide modifications ("Phosphorus Linkage" or "P Link"): -P(S), -P(Shydroxymethylphosphonate, phosphonocarboxylate), -thiophosponoacetate, thiophosphonocarboxylate, -P(CH₃)₂ (methylphosphonate, alkylphosphonate), -P(BH₃)₂ (baranophosphonate), -P(S)₂ (phosphordithioate)

Base modifications: 2-thiouracil, 6-thioguanine, 4-thiouracil, 6-thiouracil, 2-aminopurine, pseudouracil, hypoxanthine, 7-deazaguanine, 7-deaza-8-azaguanine, 5-hydroxymethyluracil, 5,6-dihydrouracil, 5-propynylcytosine, 5-ethynyluracil, 5-allyluracil, 5-aminoallyl, 5-alkylcytosine, 5-aminoallyl-uracil, 5-aminoallyl-cytosine and abasic nucleotides.

Base Pair modifications: Zfp nucleotides, UNA, isoC/isoG, 6-thioG/5-methyl-pyrimidine, x{A,G,C,T} and y{A,G,C,T}.

Other: End modifications and/or spacer/linker (ends or internal) modifications: PEG, hydrocarbon spacer, (including: heteroatom (O,S,N)-substituted hydrocarbon spacers, halo substituted hydrocarbon spacers, fluorescein-phosphoramidite and the like, squarate conjugation, Diels-Alder conjugation, or "Click" chemistry conjugation, dyes linkers including: 6-

FIG. 7

U.S. Patent

Jul. 2, 2019

Sheet 10 of 17

US 10,337,001 B2

Fluorophore-modified CLTA1 sgRNAs for in vitro testing

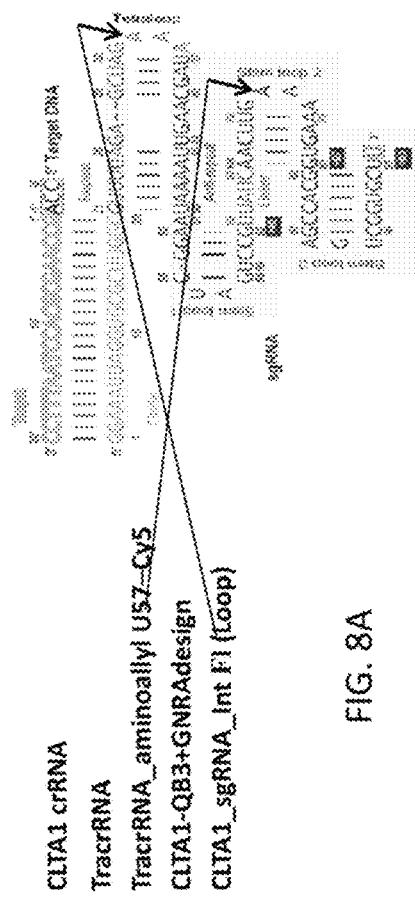


FIG. 8A

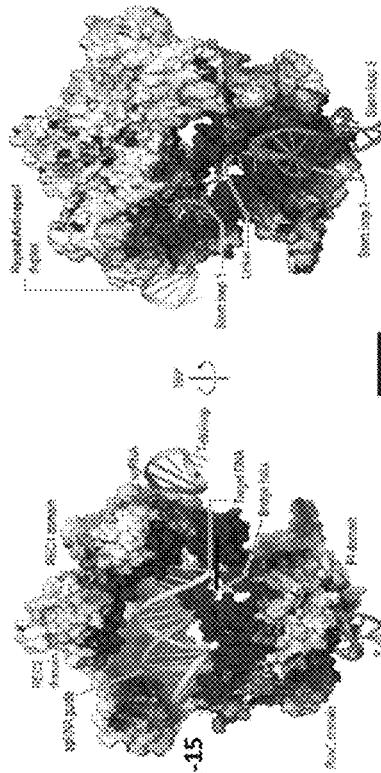


FIG. 8B

X-ray structure of Cas9 complex:
Nishimasu et al., *Cell* 2014, 156, 1-15

U.S. Patent

Jul. 2, 2019

Sheet 11 of 17

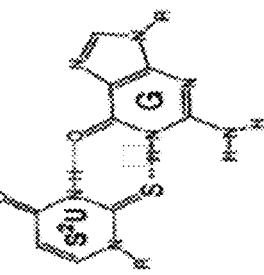
US 10,337,001 B2

Using chemical modifications to improve specificity

Target Name	CITAI On- or Off- Target Site	Genomic Coordinates	COSMIC Score	MIT Design Score
CITA1 ON1	AGTCCTCACTCCCTCAAGCAGG TCAGGAGTAGGGAGTTCGTCC	Chr9:36211735-36211757	0	100.0
CITA1 OFF1	AGTCCTCAACTCCCTCAASCAGG TCAGGAGTTGGGAGTTCGTCC	Chr8:15688928-15688950	0.35	61.1
CITA1 OFF2	AGCCCTCAATTCCCTCAAGCAGG TCGGGAGTAAGGGAGTTCGTCC	Chr3:541890184-54189106	0.65	6.4
CITA1 OFF3	ACTCCTCAACCCCCCAAGCGG TGAGGAGTAGGGAGTTCGGCC	Chr15:88845439-88845461	0.83	4.5

Example:

- CITA1_2thioU+11 crRNA: 5' AGUCCUCUAUC (2'sU) CCCUCAAAGCUUAAAGAGCUAUGCUGUUUGA AUGGUCCCCAAC 3'
- CITA1_2thioU+9
- CITA1_2thioU+3



Sulfur in 2-thioU cannot H-bond with G to form a wobble pair that lowers specificity.

FIG. 9A

U.S. Patent

Jul. 2, 2019

Sheet 12 of 17

US 10,337,001 B2

2-thioU can increase target specificity of guide RNAs when off-target sites involve U-G wobble pairing

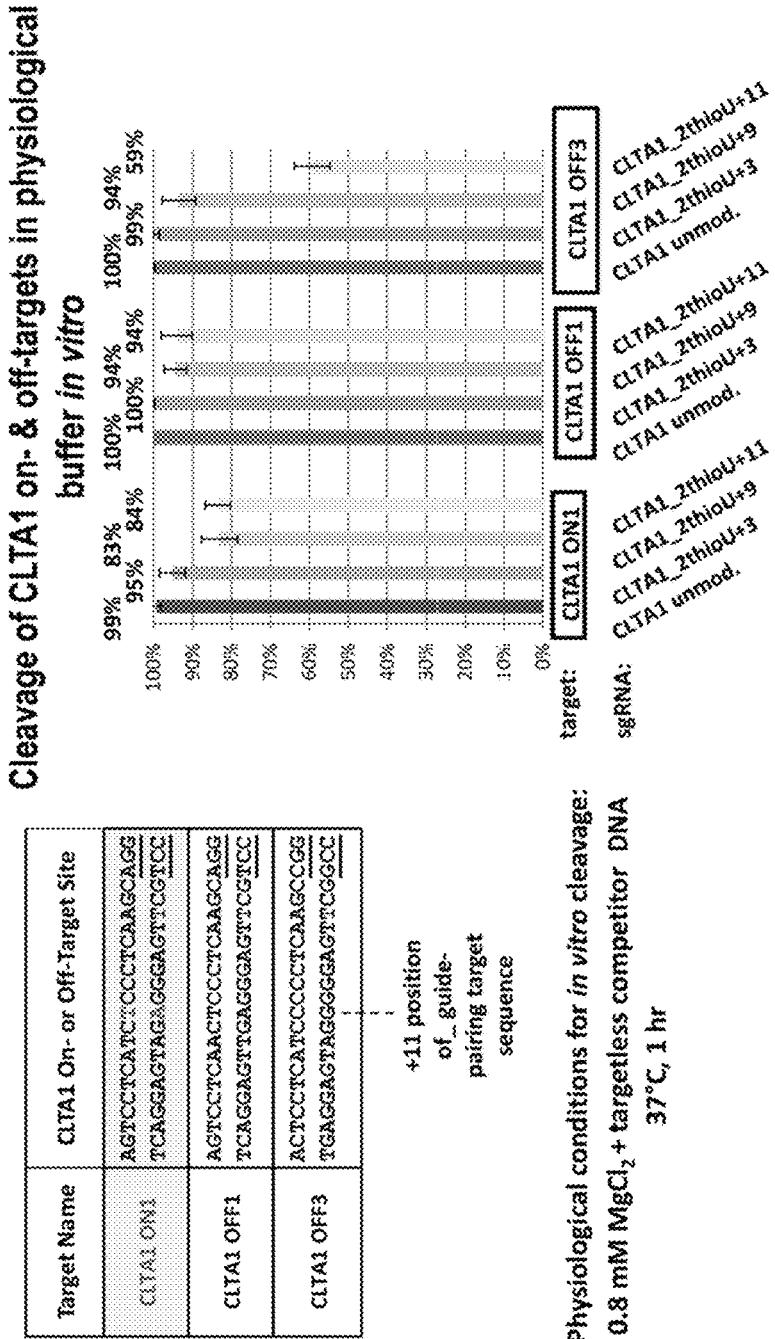


FIG. 9B

U.S. Patent

Jul. 2, 2019

Sheet 13 of 17

US 10,337,001 B2

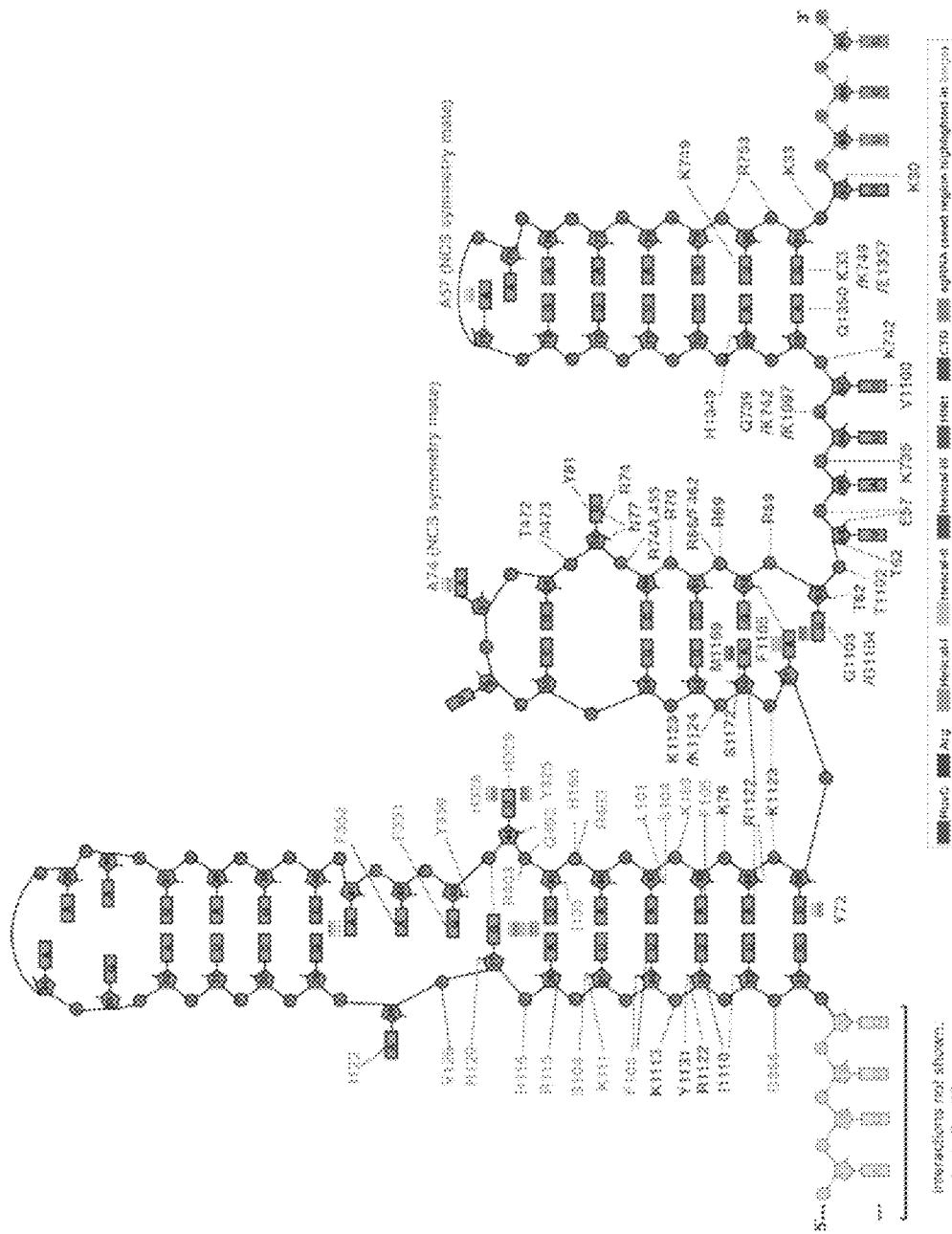


FIG. 10

U.S. Patent

Jul. 2, 2019

Sheet 14 of 17

US 10,337,001 B2

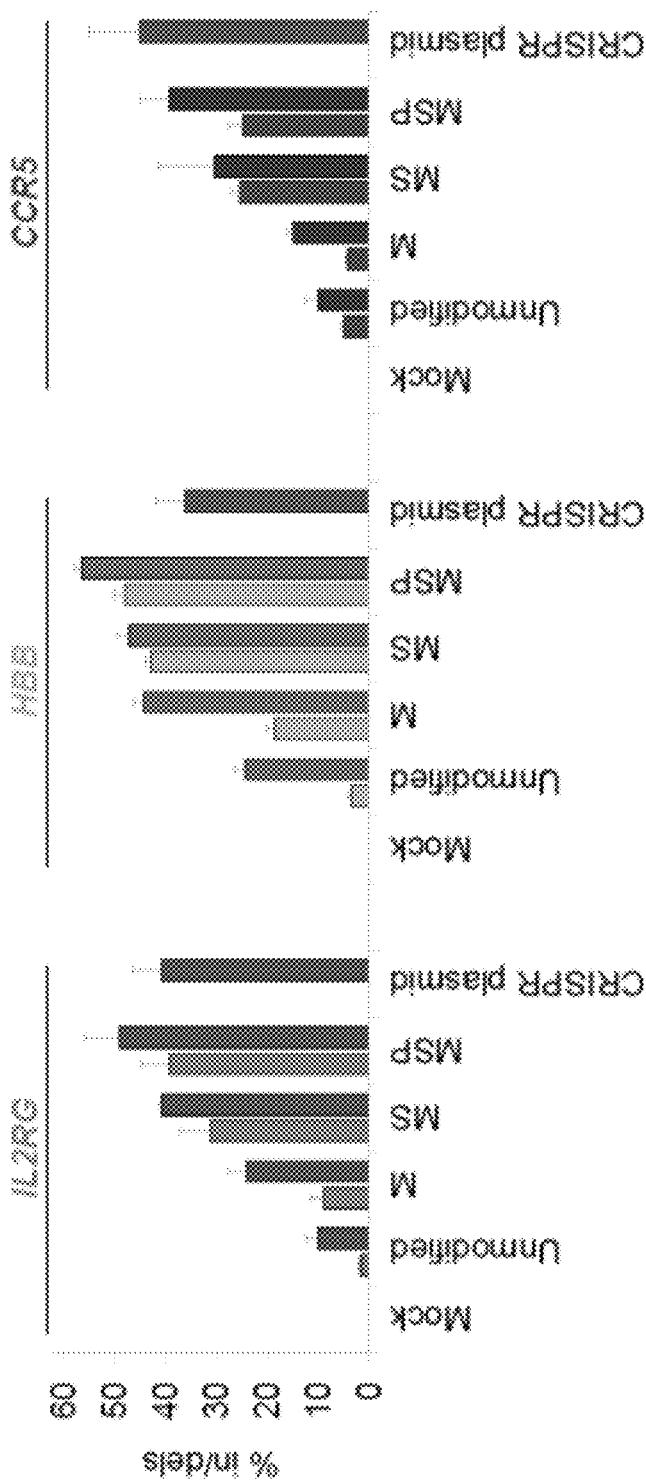


FIG. 11A

U.S. Patent

Jul. 2, 2019

Sheet 15 of 17

US 10,337,001 B2

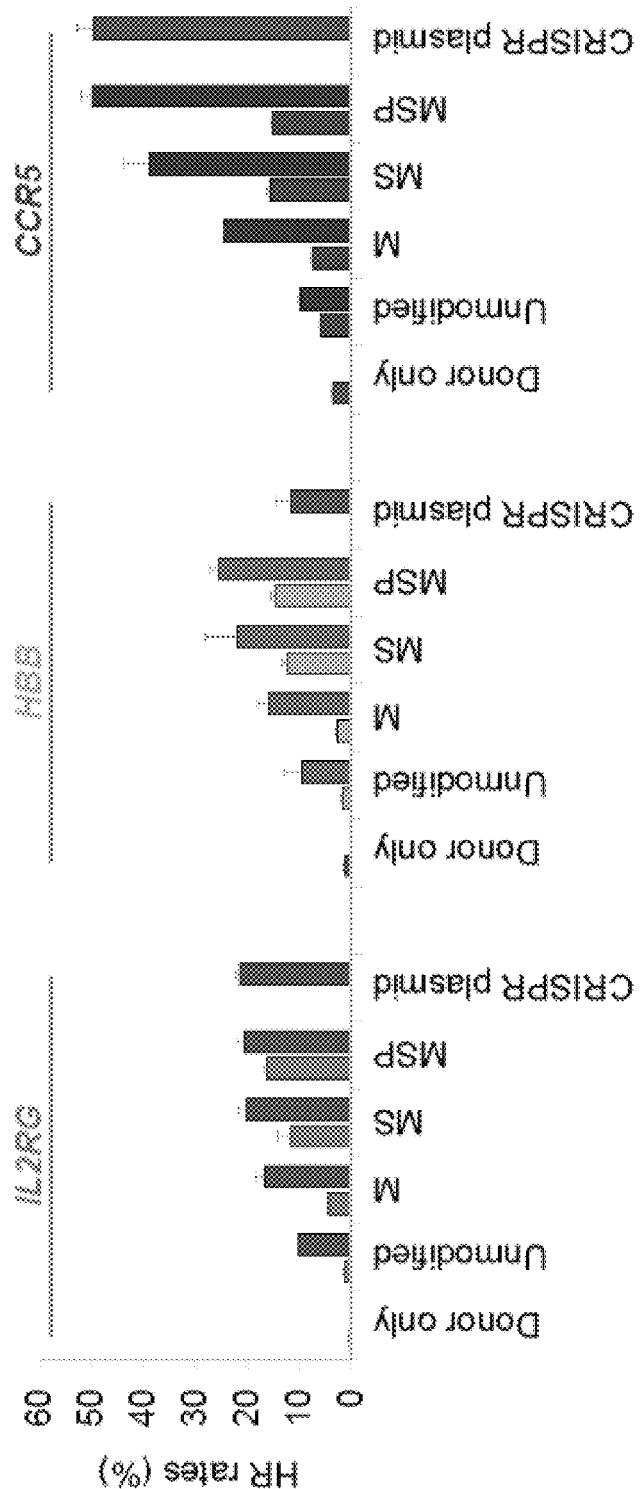


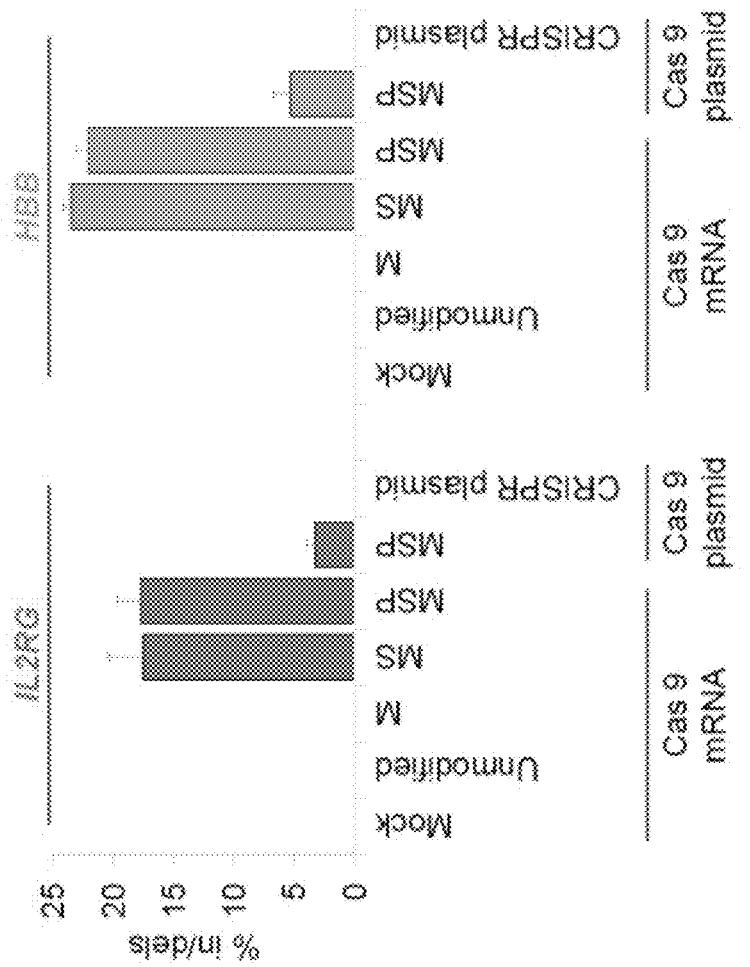
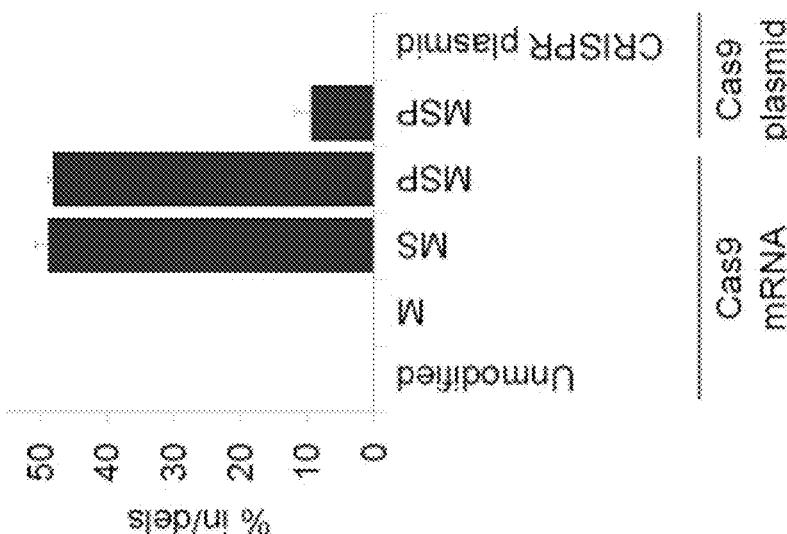
FIG. 11B

U.S. Patent

Jul. 2, 2019

Sheet 16 of 17

US 10,337,001 B2

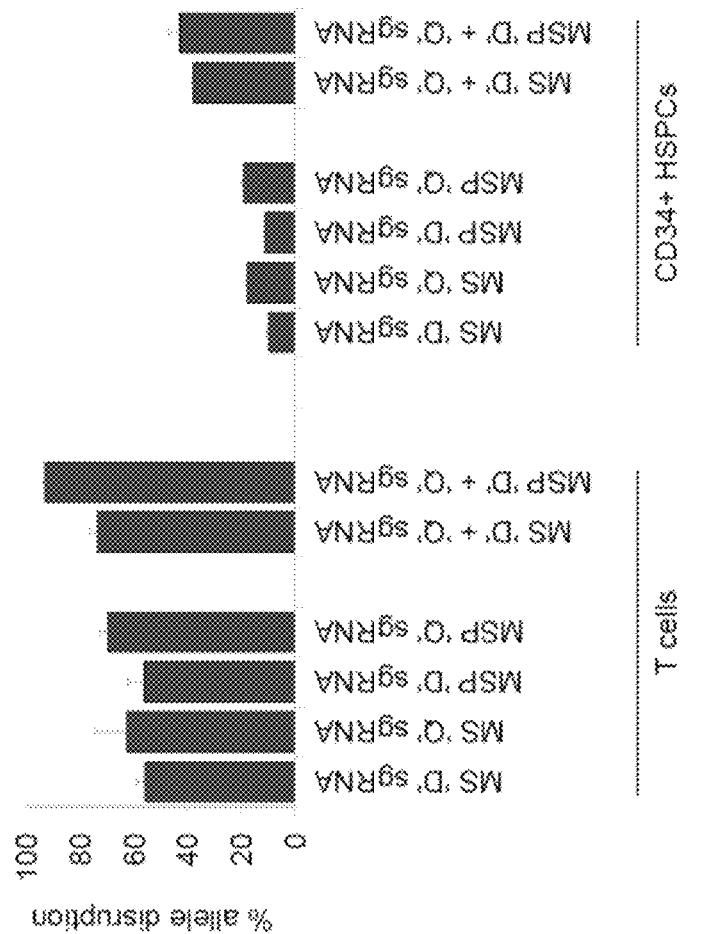
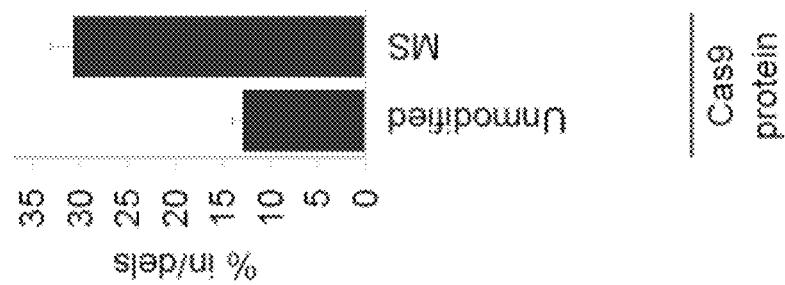
**FIG. 12C****FIG. 12A**

U.S. Patent

Jul. 2, 2019

Sheet 17 of 17

US 10,337,001 B2

**FIG. 12D****FIG. 12B**

US 10,337,001 B2

1**GUIDE RNA WITH CHEMICAL MODIFICATIONS****CROSS-REFERENCING**

This application is a continuation of U.S. application Ser. No. 14/757,204, filed on Dec. 3, 2015, and claims the benefit of U.S. Provisional Application No. 62/256,095, filed on Nov. 16, 2015; U.S. Provisional Application No. 62/146,189, filed on Apr. 10, 2015; and U.S. Provisional Application No. 62/087,211, filed on Dec. 3, 2014, all of which are incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to the field of molecular biology. In particular, the present invention relates to the clusters of regularly interspaced short palindromic repeats (CRISPR) technology.

BACKGROUND OF THE INVENTION

The native prokaryotic CRISPR-Cas system comprises an array of short repeats with intervening variable sequences of constant length (i.e., clusters of regularly interspaced short palindromic repeats, or "CRISPR"), and CRISPR-associated ("Cas") proteins. The RNA of the transcribed CRISPR array is processed by a subset of the Cas proteins into small guide RNAs, which generally have two components as discussed below. There are at least three different systems: Type I, Type II and Type III. The enzymes involved in the processing of the RNA into mature crRNA are different in the 3 systems. In the native prokaryotic system, the guide RNA ("gRNA") comprises two short, non-coding RNA species referred to as CRISPR RNA ("crRNA") and trans-acting RNA ("tracrRNA"). In an exemplary system, the gRNA forms a complex with a Cas nuclease. The gRNA:Cas nuclease complex binds a target polynucleotide sequence having a protospacer adjacent motif ("PAM") and a protospacer, which is a sequence complementary to a portion of the gRNA. The recognition and binding of the target polynucleotide by the gRNA:Cas nuclease complex induces cleavage of the target polynucleotide. The native CRISPR-Cas system functions as an immune system in prokaryotes, where gRNA:Cas nuclease complexes recognize and silence exogenous genetic elements in a manner analogous to RNAi in eukaryotic organisms, thereby conferring resistance to exogenous genetic elements such as plasmids and phages.

It has been demonstrated that a single-guide RNA ("sgRNA") can replace the complex formed between the naturally-existing crRNA and tracrRNA.

Considerations relevant to developing a gRNA, including a sgRNA, include specificity, stability, and functionality. Specificity refers to the ability of a particular gRNA:Cas nuclease complex to bind to and/or cleave a desired target sequence, whereas little or no binding and/or cleavage of polynucleotides different in sequence and/or location from the desired target occurs. Thus, specificity refers to minimizing off-target effects of the gRNA:Cas nuclease complex. Stability refers to the ability of the gRNA to resist degradation by enzymes, such as nucleases, and other substances that exist in intracellular and extra-cellular environments. Thus, there is a need for providing gRNA, including sgRNA, having increased resistance to nucleolytic degradation, increased binding affinity for the target polynucleotide, and/or reduced off-target effects while, nonetheless, having gRNA functionality. Further considerations relevant to

2

developing a gRNA include transferability and immunostimulatory properties. Thus, there is a need for providing gRNA, including sgRNA, having efficient and titratable transferability into cells, especially into the nuclei of eukaryotic cells, and having minimal or no immunostimulatory properties in the transfected cells. Another important consideration for gRNA is to provide an effective means for delivering it into and maintaining it in the intended cell, tissue, bodily fluid or organism for a duration sufficient to allow the desired gRNA functionality.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a set of diagrams showing a schematic model of an exemplary CRISPR-Cas system. The exemplary system shown here is a Type II system having a Cas nuclease. In this particular example, the Cas nuclease is the Cas9 nuclease. The Cas9 nuclease recognizes a PAM sequence (here, the PAM sequence is a 3-nt sequence of NGG, where N is A, G, C or T, but other PAM sequences are known to exist). The sgRNA includes a guide sequence, a crRNA sequence or segment, and tracrRNA sequence or segment. The guide sequence of the sgRNA hybridizes with the DNA target directly upstream of the PAM sequence. In the example shown here, Cas9 mediates a double-stranded break upstream of the PAM sequence (arrows).

FIG. 2A is a diagram showing an exemplary CRISPR-Cas9-mediated cleavage assay.

FIG. 2B is a table showing components and their concentrations for a biochemical cleavage assay used to generate the data in **FIG. 4**.

FIG. 2C is a diagram showing titration of *Streptococcus pyogenes* Cas9 nuclease for the biochemical cleavage assay.

FIG. 2D is a diagram showing titration of an exemplary sgRNA for the biochemical cleavage assay. In this example a sgRNA named kanC1 is targeted to a complementary sequence in the kanamycin resistance gene.

FIG. 3 shows exemplary conditions and procedures for the biochemical cleavage assay which uses purified components in vitro.

FIG. 4 is a table showing the data obtained using exemplary modified guide RNAs in the cleavage assay.

FIG. 5A shows an exemplary guide RNA disclosed in the application.

FIG. 5B shows an exemplary single guide RNA (sgRNA) disclosed in the application.

FIG. 6 is a table showing exemplary guide RNAs having at least two chemical modifications (e.g., a first modification and a second modification). Each number represents a modification as indicated and each "x" indicates the combination of modifications in a guide RNA. In certain embodiments, the first and second modifications are present on a single nucleotide. In certain embodiments, the first and second modifications are present on separate nucleotides.

FIG. 7 shows exemplary types of guide RNAs having at least three chemical modifications. The lower part of **FIG. 7** lists several types of modifications. The table in the upper part of **FIG. 7** indicates how a double modification ("double mod," a combination of two types of modifications) can be combined with a single modification ("single mod," one type of modification). An "x" indicates the presence of the corresponding double mod and single mod in a guide RNA.

FIGS. 8A and 8B show fluorophore-modified CLTA1 sgRNAs for in vitro testing. In **FIG. 8A**, the RNA sequence of a sgRNA for CLTA1 is shown, including a position where a fluorescent dye or a label could be attached to the sgRNA.

US 10,337,001 B2

3

FIG. 8B shows a structure determined by X-ray crystallography of a Cas9:sgRNA complex, as reported in Nishimasu et al., *Cell* (2014) 156, 1-15.

FIG. 9A shows CTLA1 sgRNAs modified with 2-thiouridine at certain locations (positions 3, 9 and 11) in an effort to improve specificity for the CTLA1 target. FIG. 9B shows that gRNA modified with 2-thioU can increase target specificity of the gRNAs when off-target sites involve U-G wobble pairing. In particular, the CTLA1_2-thioU+11 had much lower cleavage of the off-target sequence CTLA1 OFF3, which has a T to C mutation at the 11th position in the 5' strand.

FIG. 10 shows the guide RNA scaffold secondary structure, displaying noncovalent binding interactions with amino acids of Cas9, as reported in Jiang et al., *Science* (2015) 348:6242, 1477-81.

FIGS. 11A and 11B illustrate experimental results showing that gene disruption in human cell lines, with high frequencies of indels and homologous recombination (HR), can be achieved using synthesized and chemically modified sgRNAs disclosed herein, as reported in Hendel et al., *Nat. Biotechnol.* (2015) 33:9, 985-9.

FIGS. 12A, 12B, 12C and 12D illustrate experimental results showing that chemically modified sgRNAs as described herein can be used to achieve high frequencies of gene disruption or targeted genome editing in stimulated primary human T cells as well as in CD34+ hematopoietic stem and progenitor cells (HSPCs), as reported in Hendel et al., *Nat. Biotechnol.* (2015) 33:9, 985-9.

DETAILED DESCRIPTION OF THE INVENTION

This invention is based, at least in part, on an unexpected discovery that certain chemical modifications to gRNA are tolerated by the CRISPR-Cas system. In particular, certain chemical modifications believed to increase the stability of the gRNA, to alter the thermostability of a gRNA hybridization interaction, and/or to decrease the off-target effects of Cas:gRNA complexation do not substantially compromise the efficacy of Cas:gRNA binding to, nicking of, and/or cleavage of the target polynucleotide. Furthermore, certain chemical modifications are believed to provide gRNA, including sgRNA, having efficient and titratable transferability into cells, especially into the nuclei of eukaryotic cells, and/or having minimal or no immunostimulatory properties in the transfected cells. Certain chemical modifications are believed to provide gRNA, including sgRNA, which can be effectively delivered into and maintained in the intended cell, tissue, bodily fluid or organism for a duration sufficient to allow the desired gRNA functionality.

I. Definitions

As used herein, the term “guide RNA” generally refers to an RNA molecule (or a group of RNA molecules collectively) that can bind to a Cas protein and aid in targeting, the Cas protein to a specific location within a target polynucleotide (e.g., a DNA). A guide RNA can comprise a crRNA segment and a tracrRNA segment. As used herein, the term “crRNA” or “crRNA segment” refers to an RNA molecule or portion thereof that includes a polynucleotide-targeting guide sequence, a stem sequence, and, optionally, a 5'-overhang sequence. As used herein, the term “tracrRNA” or “tracrRNA segment” refers to an RNA molecule or portion thereof that includes a protein-binding segment (e.g., the protein-binding segment is capable of interacting with a

4

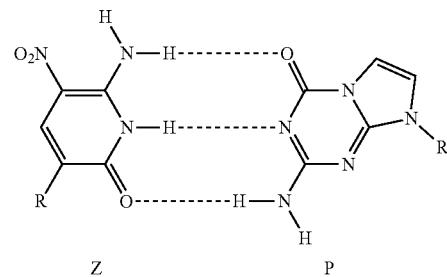
CRISPR-associated protein, such as a Cas9). The term “guide RNA” encompasses a single guide RNA (sgRNA), where the crRNA segment and the tracrRNA segment are located in the same RNA molecule. The term “guide RNA” also encompasses, collectively, a group of two or more RNA molecules, where the crRNA segment and the tracrRNA segment are located in separate RNA molecules.

The term “scaffold” refers to the portions of guide RNA molecules comprising sequences which are substantially identical or are highly conserved across natural biological species. Scaffolds include the tracrRNA segment and the portion of the crRNA segment other than the polynucleotide-targeting guide sequence at or near the 5' end of the crRNA segment, excluding any unnatural portions comprising sequences not conserved in native crRNAs and tracrRNAs.

The term “nucleic acid”, “polynucleotide” or “oligonucleotide” refers to a DNA molecule, an RNA molecule, or analogs thereof. As used herein, the terms “nucleic acid”, “polynucleotide” and “oligonucleotide” include, but are not limited to DNA molecules such as cDNA, genomic DNA or synthetic DNA and RNA molecules such as a guide RNA, messenger RNA or synthetic RNA. Moreover, as used herein, the terms “nucleic acid” and “polynucleotide” include single-stranded and double-stranded forms.

The term “modification” in the context of an oligonucleotide or polynucleotide includes but is not limited to (a) end modifications, e.g., 5' end modifications or 3' end modifications, (b) nucleobase (or “base”) modifications, including replacement or removal of bases, (c) sugar modifications, including modifications at the 2', 3', and/or 4' positions, and (d) backbone modifications, including modification or replacement of the phosphodiester linkages. The term “modified nucleotide” generally refers to a nucleotide having a modification to the chemical structure of one or more of the base, the sugar, and the phosphodiester linkage or backbone portions, including nucleotide phosphates.

The terms “Z” and “P” refer to the nucleotides, nucleobases, or nucleobase analogs developed by Steven Benner and colleagues as described for example in “Artificially expanded genetic information system: a new base pair with an alternative hydrogen bonding pattern” Yang, Z., Hutter, D., Sheng, P., Sismour, A. M. and Benner, S. A. (2006) *Nucleic Acids Res.*, 34, 6095-101, the contents of which is hereby incorporated by reference in its entirety.



The terms “xA”, “xG”, “xC”, “xT”, or “x(A,G,C,T)” and “yA”, “yG”, “yC”, “yT”, or “y(A,G,C,T)” refer to nucleotides, nucleobases, or nucleobase analogs as described by Krueger et al in “Synthesis and Properties of Size-Expanded DNAs: Toward Designed, Functional Genetic Systems”; Andrew T. Krueger, Haige Lu, Alex H. F. Lee, and Eric T. Kool (2007) *Acc. Chem. Res.*, 40, 141-50, the contents of which is hereby incorporated by reference in its entirety.

US 10,337,001 B2

5

The term “Unstructured Nucleic Acid” or “UNA” refers to nucleotides, nucleobases, or nucleobase analogs as described in U.S. Pat. No. 7,371,580, the contents of which is hereby incorporated by reference in its entirety. An unstructured nucleic acid, or UNA, modification is also referred to as a “pseudo-complementary” nucleotide, nucleobase or nucleobase analog (see e.g., Lahoud et al. (1991) *Nucl. Acids Res.*, 36:10, 3409-19).

The terms “PACE” and “thioPACE” refer to internucleotide phosphodiester linkage analogs containing phosphonoacetate or thiophosphonoacetate groups, respectively. These modifications belong to a broad class of compounds comprising phosphonocarboxylate moiety, phosphonocarboxylate ester moiety, thiophosphonocarboxylate moiety and thiophosphonocarboxylate ester moiety. These linkages can be described respectively by the general formulae P(CR₁R₂)_nCOOR and (S)—P(CR₁R₂)_nCOOR wherein n is an integer from 0 to 6 and each of R₁ and R₂ is independently selected from the group consisting of H, an alkyl and substituted alkyl. Some of these modifications are described by Yamada et al. in “Synthesis and Biochemical Evaluation of Phosphonoformate Oligodeoxyribonucleotides” Christina M. Yamada, Douglas J. Dellinger and Marvin H. Caruthers (2006) *J. Am. Chem. Soc.* 128:15, 5251-61, the contents of which is hereby incorporated by reference in its entirety.

As used herein, “modification” refers to a chemical moiety, or portion of a chemical structure, which differs from that found in unmodified ribonucleotides, namely adenosine, guanosine, cytidine, and uridine ribonucleotides. The term “modification” may refer to type of modification. For example, “same modification” means same type of modification, and “the modified nucleotides are the same” means the modified nucleotides have the same type(s) of modification while the base (A, G, C, U, etc.) may be different. Similarly, a guide RNA with “two modifications” is a guide RNA with two types of modifications, which may or may not be in the same nucleotide, and each type may appear in multiple nucleotides in the guide RNA. Similarly, a guide RNA with “three modifications” is a guide RNA with three types of modifications, which may or may not be in the same nucleotide, and each type may appear in multiple nucleotides.

As used herein, the term “target polynucleotide” or “target” refers to a polynucleotide containing a target nucleic acid sequence. A target polynucleotide may be single-stranded or double-stranded, and, in certain embodiments, is double-stranded DNA. In certain embodiments, the target polynucleotide is single-stranded RNA. A “target nucleic acid sequence” or “target sequence,” as used herein, means a specific sequence or the complement thereof that one wishes to bind to, nick, or cleave using a CRISPR system.

The term “hybridization” or “hybridizing” refers to a process where completely or partially complementary polynucleotide strands come together under suitable hybridization conditions to form a double-stranded structure or region in which the two constituent strands are joined by hydrogen bonds. As used herein, the term “partial hybridization” includes where the double-stranded structure or region contains one or more bulges or mismatches. Although hydrogen bonds typically form between adenine and thymine or adenine and uracil (A and T or A and U) or cytosine and guanine (C and G), other noncanonical base pairs may form (See e.g., Adams et al., “The Biochemistry of the Nucleic Acids,” 11th ed., 1992). It is contemplated that modified nucleotides may form hydrogen bonds that allow or promote hybridization.

6

The term “cleavage” or “cleaving” refers to breaking of the covalent phosphodiester linkage in the ribosylphosphodiester backbone of a polynucleotide. The terms “cleavage” or “cleaving” encompass both single-stranded breaks and double-stranded breaks. Double-stranded cleavage can occur as a result of two distinct single-stranded cleavage events. Cleavage can result in the production of either blunt ends or staggered ends.

The term “CRISPR-associated protein” or “Cas protein” refers to a wild type Cas protein, a fragment thereof, or a mutant or variant thereof. The term “Cas mutant” or “Cas variant” refers to a protein or polypeptide derivative of a wild type Cas protein, e.g., a protein having one or more point mutations, insertions, deletions, truncations, a fusion protein, or a combination thereof. In certain embodiments, the “Cas mutant” or “Cas variant” substantially retains the nuclease activity of the Cas protein. In certain embodiments, the “Cas mutant” or “Cas variant” is mutated such that one or both nuclease domains are inactive. In certain embodiments, the “Cas mutant” or “Cas variant” has nuclease activity. In certain embodiments, the “Cas mutant” or “Cas variant” lacks some or all of the nuclease activity of its wild-type counterpart.

The term “nuclease domain” of a Cas protein refers to the polypeptide sequence or domain within the protein which possesses the catalytic activity for DNA cleavage. A nuclease domain can be contained in a single polypeptide chain, or cleavage activity can result from the association of two (or more), polypeptides. A single nuclease domain may consist of more than one isolated stretch of amino acids within a given polypeptide. Examples of these domains include RuvC-like motifs (amino acids 7-22, 759-766 and 982-989 in SEQ ID NO: 1) and HNH motif (aa 837-863). Sec Gasiunas et al. (2012) *Proc. Natl. Acad. Sci. USA*, 109:39, E2579-E2586 and WO2013176772.

A synthetic guide RNA that has “gRNA functionality” is one that has one or more of the functions of naturally occurring guide RNA, such as associating with a Cas protein, or a function performed by the guide RNA in association with a Cas protein. In certain embodiments, the functionality includes binding a target polynucleotide. In certain embodiments, the functionality includes targeting a Cas protein or a gRNA:Cas protein complex to a target polynucleotide. In certain embodiments, the functionality includes nicking a target polynucleotide. In certain embodiments, the functionality includes cleaving a target polynucleotide. In certain embodiments, the functionality includes associating with or binding to a Cas protein. In certain embodiments, the functionality is any other known function of a guide RNA in a CRISPR-Cas system with a Cas protein, including an artificial CRISPR-Cas system with an engineered Cas protein. In certain embodiments, the functionality is any other function of natural guide RNA. The synthetic guide RNA may have gRNA functionality to a greater or lesser extent than a naturally occurring guide RNA. In certain embodiments, a synthetic guide RNA may have greater functionality as to one property and lesser functionality as to another property in comparison to a similar naturally occurring guide RNA.

A “Cas protein having a single-strand nicking activity” refers to a Cas protein, including a Cas mutant or Cas variant, that has reduced ability to cleave one of two strands of a dsDNA as compared to a wild type Cas protein. For example, in certain embodiments, a Cas protein having a single-strand nicking activity has a mutation (e.g., amino acid substitution) that reduces the function of the RuvC domain (or the HNH domain) and as a result reduces the

US 10,337,001 B2

7

ability to cleave one strand of the target DNA. Examples of such variants include the D10A, H839A/H840A, and/or N863A substitutions in *S. pyogenes* Cas9, and also include the same or similar substitutions at equivalent sites in Cas9 enzymes of other species.

As used herein, the term “portion” or “fragment” of a sequence refers to any portion of the sequence (e.g., a nucleotide subsequence or an amino acid subsequence) that is smaller than the complete sequence. Portions of polynucleotides can be any length, for example, at least 5, 10, 15, 20, 25, 30, 40, 50, 75, 100, 150, 200, 300 or 500 or more nucleotides in length. A portion of a guide sequence can be about 50%, 40%, 30%, 20%, 10% of the guide sequence, e.g., one-third of the guide sequence or shorter, e.g., 7, 6, 5, 4, 3, or 2 nucleotides in length.

The term “derived from” in the context of a molecule refers to a molecule isolated or made using a parent molecule or information from that parent molecule. For example, a Cas9 single mutant nuclease and a Cas9 double mutant null-nuclease are derived from a wild-type Cas9 protein.

The term “substantially identical” in the context of two or more polynucleotides (or two or more polypeptides) refers to sequences or subsequences that have at least about 60%, at least about 70%, at least about 80%, at least about 90%, about 90-95%, at least about 95%, at least about 98%, at least about 99% or more nucleotide (or amino acid) sequence identity, when compared and aligned for maximum correspondence using a sequence comparison algorithm or by visual inspection. Preferably, the “substantial identity” between polynucleotides exists over a region of the polynucleotide at least about 50 nucleotides in length, at least about 100 nucleotides in length, at least about 200 nucleotides in length, at least about 300 nucleotides in length, at least about 500 nucleotides in length, or over the entire length of the polynucleotide. Preferably, the “substantial identity” between polypeptides exists over a region of the polypeptide at least about 50 amino acid residues in length, at least about 100 amino acid residues in length, or over the entire length of the polypeptide.

As disclosed herein, a number of ranges of values are provided. It is understood that each intervening value, to the tenth of the unit of the lower limit, between the upper and lower limits of that range is also specifically contemplated. Each smaller range or intervening value encompassed by a stated range is also specifically contemplated. The term “about” generally refers to plus or minus 10% of the indicated number. For example, “about 10%” may indicate a range of 9% to 11%, and “about 20” may mean from 18-22. Other meanings of “about” may be apparent from the context, such as rounding off, so, for example “about 1” may also mean from 0.5 to 1.4.

II. CRISPR-Mediated Sequence-Specific Binding and/or Cleavage

Shown in FIG. 1 is a diagram of CRISPR-Cas9-mediated sequence-specific cleavage of DNA. The guide RNA is depicted as sgRNA with an exemplary 20-nucleotide (20-nt) guide sequence (other guide sequences may be, for example, from about 15 to about 30 nts in length) within the 5' domain, an internally positioned base-paired stem, and a 3' domain. The guide sequence is complementary to an exemplary 20-nt target sequence in a DNA target. The stem corresponds to a repeat sequence in crRNA and is complementary to a sequence in the tracrRNA. The 3' domain of the guide RNA corresponds to the 3' domain of the tracrRNA

8

that binds a Cas9 nuclease. The Cas9:guide RNA complex binds and cleaves a target DNA sequence or protospacer directly upstream of a PAM sequence recognized by Cas9. In FIG. 1, a 3-nt PAM sequence is exemplified; however other PAM sequences, including 4-nt and 5-nt PAM sequences are known. In the system exemplified in FIG. 1, both strands of the target sequence in DNA are cleaved by Cas9 at the sites indicated by arrows.

10

III. Guide RNAs

In at least one aspect, the present invention comprises a chemically modified guide RNA that has guide RNA functionality. A guide RNA that comprises any nucleotide other than the four canonical ribonucleotides, namely A, C, G, and U, whether unnatural or natural (e.g., a pseudouridine, inosine or a deoxynucleotide), is a chemically modified guide RNA. Likewise a guide RNA that comprises any backbone or internucleotide linkage other than a natural phosphodiester internucleotide linkage possesses a chemical modification and therefore is a chemically modified guide RNA. In certain embodiments, the retained functionality includes binding a Cas protein. In certain embodiments, the retained functionality includes binding a target polynucleotide. In certain embodiments, the retained functionality includes targeting a Cas protein or a gRNA:Cas protein complex to a target polynucleotide. In certain embodiments, the retained functionality includes nicking a target polynucleotide by a gRNA:Cas protein complex. In certain embodiments, the retained functionality includes cleaving a target polynucleotide by a gRNA:Cas protein complex. In certain embodiments, the retained functionality is any other known function of a guide RNA in a CRISPR-Cas system with a Cas protein, including an artificial CRISPR-Cas system with an engineered Cas protein. In certain embodiments, the retained functionality is any other function of a natural guide RNA.

40

A. Exemplary Modifications

In certain embodiments, a nucleotide sugar modification incorporated into the guide RNA is selected from the group consisting of 2'-O—C₁₋₄alkyl such as 2'-O-methyl (2'-OMe), 2'-deoxy (2'-H), 2'-O—C₁₋₃alkyl-O—C₁₋₃alkyl such as 2'-methoxyethyl (“2'-MOE”), 2'-fluoro (“2'-F”), 2'-amino (“2'-NH₂”), 2'-arabinosyl (“2'-arabino”) nucleotide, 2'-F-arabinosyl (“2'-F-arabino”) nucleotide, 2'-locked nucleic acid (“LNA”) nucleotide, 2'-unlocked nucleic acid (“ULNA”) nucleotide, a sugar in L form (“L-sugar”), and 4'-thioribosyl nucleotide. In certain embodiments, an internucleotide linkage modification incorporated into the guide RNA is selected from the group consisting of: phosphorothioate “P(S)” (P(S)), phosphonocarboxylate (P(CH₂)_nCOOR) such as phosphonoacetate “PACE” (P(CH₂COO⁻)), thiophosphonocarboxylate ((S)P(CH₂)_nCOOR) such as thiophosphonoacetate “thioPACE” ((S)P(CH₂COO⁻)), alkylphosphonate (P(C₁₋₃alkyl) such as methylphosphonate —P(CH₃), boranophosphonate (P(BH₃)), and phosphorodithioate (P(S)₂).

60

In certain embodiments, a nucleobase (“base”) modification incorporated into the guide RNA is selected from the group consisting of: 2-thiouracil (“2-thioU”), 2-thiocytosine (“2-thioC”), 4-thiouracil (“4-thioU”), 6-thioguanine (“6-thioG”), 2-aminoadenine (“2-aminoA”), 2-aminopurine, pseudouracil, hypoxanthine, 7-deazaguanine, 7-deaza-8-azaguanine, 7-deazaadenine, 7-deaza-8-azaadenine, 5-methylcytosine (“5-methylC”), 5-methyluracil (“5-methylU”),

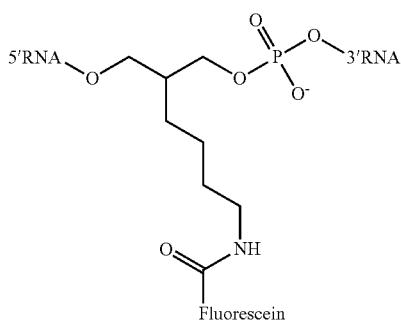
US 10,337,001 B2

9

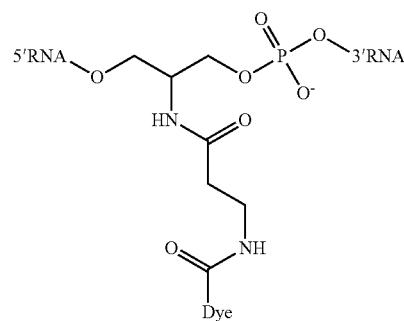
5-hydroxymethylcytosine, 5-hydroxymethyluracil, 5,6-dihydrouracil, 5-propynylcytosine, 5-propynyluracil, 5-ethynylcytosine, 5-ethynyluracil, 5-allyluracil (“5-allylU”), 5-allylcytosine (“5-allylC”), 5-aminoallyluracil (“5-aminoallylU”), 5-aminoallylcytosine (“5-aminoallylC”), an abasic nucleotide, Z base, P base, Unstructured Nucleic Acid (“UNA”), isoguanine (“isoG”), isocytosine (“isoC”) [as described in “Enzymatic Incorporation of a New Base pair into DNA and RNA Extends the Genetic Alphabet.” Piccirilli, J. A.; Krauch, T.; Moroncy, S. E.; Benner, S. A. (1990) *Nature*, 343, 33], 5-methyl-2-pyrimidine [as described in Rappaport, H. P. (1993) *Biochemistry*, 32, 3047], x(A,G,C, T) and y(A,G,C,T).

In certain embodiments, one or more isotopic modifications are introduced on the nucleotide sugar, the nucleobase, the phosphodiester linkage and/or the nucleotide phosphates. Such modifications include nucleotides comprising one or more ^{15}N , ^{13}C , ^{14}C , Deuterium, ^3H , ^{32}P , ^{125}I , ^{131}I atoms or other atoms or elements used as tracers.

In certain embodiments, an “end” modification incorporated into the guide RNA is selected from the group consisting of: PEG (polyethyleneglycol), hydrocarbon linkers (including: heteroatom (O,S,N)-substituted hydrocarbon spacers; halo-substituted hydrocarbon spacers; keto-, carboxyl-, amido-, thionyl-, carbamoyl-, thionocarbamoyl-containing hydrocarbon spacers), spermine linkers, dyes including fluorescent dyes (for example fluoresceins, rhodamines, cyanines) attached to linkers such as for example 6-fluorescein-hexyl, quenchers (for example dabcyl, BHQ) and other labels (for example biotin, digoxigenin, acridine, streptavidin, avidin, peptides and/or proteins). In certain embodiments, an “end” modification comprises a conjugation (or ligation) of the guide RNA to another molecule comprising an oligonucleotide (comprising deoxynucleotides and/or ribonucleotides), a peptide, a protein, a sugar, an oligosaccharide, a steroid, a lipid, a folic acid, a vitamin and/or other molecule. In certain embodiments, an “end” modification incorporated into the guide RNA is located internally in the guide RNA sequence via a linker such as for example 2-(4-butylamidofluorescein)propane-1,3-diol bis (phosphodiester) linker (depicted below), which is incorporated as a phosphodiester linkage and can be incorporated anywhere between two nucleotides in the guide RNA.



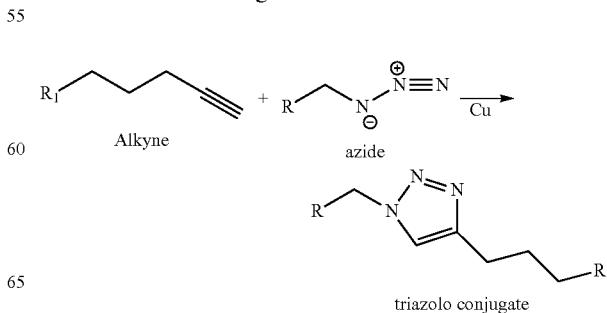
2-(4-butylamidofluorescein)propane-1,3-diol
bis(phosphodiester) Linker

10

2-(3-(dye-amido)propanamido)propane-1,3-diol bis
(phosphodiester) Linker

In certain embodiments, the end modification comprises a terminal functional group such as an amine, a thiol (or sulhydryl), a hydroxyl, a carboxyl, carbonyl, thionyl, thiocarbonyl, a carbamoyl, a thiocarbamoyl, a phosphoryl, an alkene, an alkyne, an halogen or a functional group-terminated linker, either of which can be subsequently conjugated to a desired moiety, for example a fluorescent dye or a non-fluorescent label or tag or any other molecule such as for example an oligonucleotide (comprising deoxynucleotides and/or ribonucleotides, including an aptamer), an amino acid, a peptide, a protein, a sugar, an oligosaccharide, a steroid, a lipid, a folic acid, a vitamin. The conjugation employs standard chemistry well-known in the art, including but not limited to coupling via N-hydroxysuccinimide, iso-thiocyanate, DCC (or DCI), and/or any other standard method as described in “Bioconjugate Techniques” by Greg T. Hermanson, Publisher Elsevier Science, 3rd ed. (2013), the contents of which are incorporated herein by reference in their entireties.

In certain embodiments, the label or dye is attached or conjugated to a modified nucleotide in the gRNA. The conjugation of a fluorescent dye or other moiety such as a non-fluorescent label or tag (for example biotin, avidin, streptavidin, or moiety containing an isotopic label such as ^{15}N , ^{13}C , ^{14}C , Deuterium, ^3H , ^{32}P , ^{125}I and the like) or any other molecule such as for example an oligonucleotide (comprising deoxynucleotides and/or ribonucleotides including an aptamer), an amino acid, a peptide, a protein, a sugar, an oligosaccharide, a steroid, a lipid, a folic acid, a vitamin or other molecule can be effectuated using the so-called “click” chemistry or the so-called “squarate” conjugation chemistry. The “click” chemistry refers to the [3+2] cycloaddition of an alkyne moiety with an azide moiety, leading to a triazolo linkage between the two moieties as shown in the following scheme:



Other linkers include for example by way of illustration, but are not limited to:

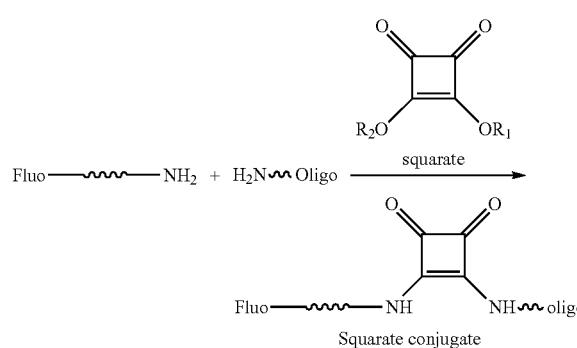
US 10,337,001 B2

11

as described for example in El-Sagheer, A. H and Brown, T. “Click chemistry with DNA”, *Chem. Soc. Rev.*, 2010, 39, 1388-1405 and Mojibul, H. M. and XiaoHua, P., DNA-associated click chemistry, *Sci. China Chem.*, 2014, 57:2, 215-31, the contents of which are hereby incorporated by reference in their entirety.

In certain embodiments, the conjugation can be effectuated by alternative cycloaddition such as Diels-Alder [4+2] cycloaddition of a π -conjugated diene moiety with an alkene moiety.

The “squarate” conjugation chemistry links two moieties each having an amine via a squarate derivative to result in a squarate conjugate that contains a squarate moiety (see e.g., Tietze et al. (1991) *Chem. Ber.*, 124, 1215-21, the contents of which are hereby incorporated by reference in their entirety). For example, a fluorescein containing a linker amine is conjugated to an oligoribonucleotide containing an amine through a squarate linker as described in the scheme below. An example of the squarate linker is depicted in the following scheme:

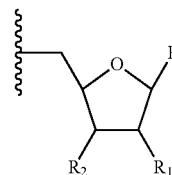


In certain, embodiments, a chemical modification incorporated into the guide RNA is selected from the group consisting of 2'-O—C₁₋₄alkyl, 2'-H, 2'-O—C₁₋₃alkyl-O—C₁₋₃alkyl, 2'-F, 2'-NH₂, 2'-arabino, 2'-F-arabino, 4'-thioribosyl, 2-thioU, 2-thioC, 4-thioU, 6-thioG, 2-aminoA, 2-aminopurine, pseudouracil, hypoxanthine, 7-deazaguanine, 7-deaza-8-azaguanine, 7-deazaadenine, 7-deaza-8-azaadenine, 5-methylC, 5-methylU, 5-hydroxymethylcytosine, 5-hydroxymethyluracil, 5,6-dehydrouracil, 5-propynylcytosine, 5-propynyluracil, 5-ethynylcytosine, 5-ethynyluracil, 5-allylU, 5-allylC, 5-aminoallyl-uracil, 5-aminoallyl-cytosine, an abasic nucleotide (“abN”), Z, P, UNA, isoC, isoG, 5-methyl-pyrimidine, x(A,G,C,T) and y(A,G,C,T), a phosphorothioate internucleotide linkage, a phosphonoacetate internucleotide linkage, a thiophosphonoacetate internucleotide linkage, a methylphosphonate internucleotide linkage, a boranophosphonate internucleotide linkage, a phosphorodithioate internucleotide linkage, 4-thioribosyl nucleotide, a locked nucleic acid (“LNA”) nucleotide, an unlocked nucleic acid (“ULNA”) nucleotide, an alkyl spacer, a heteroalkyl (N, O, S) spacer, a 5'- and/or 3'-alkyl terminated nucleotide, a Unicap, a 5'-terminal cap known from nature, an xRNA base (analogous to “xDNA” base), an yRNA base (analogous to “yDNA” base), a PEG substituent, or a conjugated linker to a dye or non-fluorescent label (or tag) or other moiety as described above. Exemplary modified nucleotides are also depicted in Table 2.

12

TABLE 2

Exemplary modified nucleotides contained in a synthetic guide sequence.



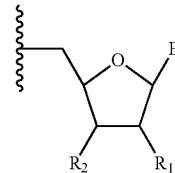
#	R ₁	R ₂	B
A1	OH	OH	uridine
A2	OMe	OH	uridine
15 A3	F	OH	uridine
A4	Cl	OH	uridine
A5	Br	OH	uridine
A6	I	OH	uridine
A7	NH ₂	OH	uridine
A8	H	OH	uridine
20 A9	OH	phosphodiester	uridine
A10	OMe	phosphodiester	uridine
A11	F	phosphodiester	uridine
A12	Cl	phosphodiester	uridine
A13	Br	phosphodiester	uridine
A14	I	phosphodiester	uridine
A15	NH ₂	phosphodiester	uridine
A16	H	phosphodiester	uridine
25 A17	OH	phosphonoacetate	uridine
A18	OMe	phosphonoacetate	uridine
A19	F	phosphonoacetate	uridine
A20	Cl	phosphonoacetate	uridine
A21	Br	phosphonoacetate	uridine
30 A22	I	phosphonoacetate	uridine
A23	NH ₂	phosphonoacetate	uridine
A24	H	phosphonoacetate	uridine
A25	OH	thiophosphonoacetate	uridine
35 A26	OMe	thiophosphonoacetate	uridine
A27	F	thiophosphonoacetate	uridine
A28	Cl	thiophosphonoacetate	uridine
A29	Br	thiophosphonoacetate	uridine
A30	I	thiophosphonoacetate	uridine
A31	NH ₂	thiophosphonoacetate	uridine
A32	H	thiophosphonoacetate	uridine
A33	OH	phosphorothioate	uridine
40 A34	OMe	phosphorothioate	uridine
A35	F	phosphorothioate	uridine
A36	Cl	phosphorothioate	uridine
A37	Br	phosphorothioate	uridine
A38	I	phosphorothioate	uridine
A39	NH ₂	phosphorothioate	uridine
A40	H	phosphorothioate	uridine
A41	OH	phosphordithioate	uridine
45 A42	OMe	phosphordithioate	uridine
A43	F	phosphordithioate	uridine
A44	Cl	phosphordithioate	uridine
A45	Br	phosphordithioate	uridine
A46	I	phosphordithioate	uridine
50 A47	NH ₂	phosphordithioate	uridine
A48	H	phosphordithioate	uridine
A49	OH	methylphosphonate	uridine
A50	OMe	methylphosphonate	uridine
A51	F	methylphosphonate	uridine
A52	Cl	methylphosphonate	uridine
55 A53	Br	methylphosphonate	uridine
A54	I	methylphosphonate	uridine
A55	NH ₂	methylphosphonate	uridine
A56	H	methylphosphonate	uridine
A57	OH	boranophosphonate	uridine
60 A58	OMe	boranophosphonate	uridine
A59	F	boranophosphonate	uridine
A60	Cl	boranophosphonate	uridine
A61	Br	boranophosphonate	uridine
A62	I	boranophosphonate	uridine
65 A63	NH ₂	boranophosphonate	uridine
A64	H	boranophosphonate	uridine
B1	OH	OH	adenosine
B2	OMe	OH	adenosine

US 10,337,001 B2

13

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.

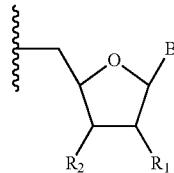


#	R ₁	R ₂	B
B3	F	OH	adenosine
B4	Cl	OH	adenosine
B5	Br	OH	adenosine
B6	I	OH	adenosine
B7	NH ₂	OH	adenosine
B8	H	OH	adenosine
B9	OH	phosphodiester	adenosine
B10	OMe	phosphodiester	adenosine
B11	F	phosphodiester	adenosine
B12	Cl	phosphodiester	adenosine
B13	Br	phosphodiester	adenosine
B14	I	phosphodiester	adenosine
B15	NH ₂	phosphodiester	adenosine
B16	H	phosphodiester	adenosine
B17	OH	phosphonoacetate	adenosine
B18	OMe	phosphonoacetate	adenosine
B19	F	phosphonoacetate	adenosine
B20	Cl	phosphonoacetate	adenosine
B21	Br	phosphonoacetate	adenosine
B22	I	phosphonoacetate	adenosine
B23	NH ₂	phosphonoacetate	adenosine
B24	H	phosphonoacetate	adenosine
B25	OH	thiophosphonoacetate	adenosine
B26	OMe	thiophosphonoacetate	adenosine
B27	F	thiophosphonoacetate	adenosine
B28	Cl	thiophosphonoacetate	adenosine
B29	Br	thiophosphonoacetate	adenosine
B30	I	thiophosphonoacetate	adenosine
B31	NH ₂	thiophosphonoacetate	adenosine
B32	H	thiophosphonoacetate	adenosine
B33	OH	phosphorothioate	adenosine
B34	OMe	phosphorothioate	adenosine
B35	F	phosphorothioate	adenosine
B36	Cl	phosphorothioate	adenosine
B37	Br	phosphorothioate	adenosine
B38	I	phosphorothioate	adenosine
B39	NH ₂	phosphorothioate	adenosine
B40	H	phosphorothioate	adenosine
B41	OH	phosphorodithioate	adenosine
B42	OMe	phosphorodithioate	adenosine
B43	F	phosphorodithioate	adenosine
B44	Cl	phosphorodithioate	adenosine
B45	Br	phosphorodithioate	adenosine
B46	I	phosphorodithioate	adenosine
B47	NH ₂	phosphorodithioate	adenosine
B48	H	phosphorodithioate	adenosine
B49	OH	methylphosphonate	adenosine
B50	OMe	methylphosphonate	adenosine
B51	F	methylphosphonate	adenosine
B52	Cl	methylphosphonate	adenosine
B53	Br	methylphosphonate	adenosine
B54	I	methylphosphonate	adenosine
B55	NH ₂	methylphosphonate	adenosine
B56	H	methylphosphonate	adenosine
B57	OH	boranophosphonate	adenosine
B58	OMe	boranophosphonate	adenosine
B59	F	boranophosphonate	adenosine
B60	Cl	boranophosphonate	adenosine
B61	Br	boranophosphonate	adenosine
B62	I	boranophosphonate	adenosine
B63	NH ₂	boranophosphonate	adenosine
B64	H	boranophosphonate	adenosine
C1	OH	cytidine	cytidine
C2	OMe	OH	cytidine
C3	F	OH	cytidine
C4	Cl	OH	cytidine

14

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.



5			
10			
15	C7	NH ₂	OH
16	C8	H	OH
17	C9	OH	phosphodiester
18	C10	OMe	phosphodiester
19	C11	F	phosphodiester
20	C12	Cl	phosphodiester
21	C13	Br	phosphodiester
22	C14	I	phosphodiester
23	C15	NH ₂	phosphodiester
24	C16	H	phosphodiester
25	C17	OH	phosphonoacetate
26	C18	OMe	phosphonoacetate
27	C19	F	phosphonoacetate
28	C20	Cl	phosphonoacetate
29	C21	Br	phosphonoacetate
30	C22	I	phosphonoacetate
31	C23	NH ₂	phosphonoacetate
32	C24	H	phosphonoacetate
33	C25	OH	thiophosphonoacetate
34	C26	OMe	thiophosphonoacetate
35	C27	F	thiophosphonoacetate
36	C28	Cl	thiophosphonoacetate
37	C29	Br	thiophosphonoacetate
38	C30	I	thiophosphonoacetate
39	C31	NH ₂	thiophosphonoacetate
40	C32	H	thiophosphonoacetate
41	C33	OH	phosphorothioate
42	C34	OMe	phosphorothioate
43	C35	F	phosphorothioate
44	C36	Cl	phosphorothioate
45	C37	Br	phosphorothioate
46	C38	I	phosphorothioate
47	C39	NH ₂	phosphorothioate
48	C40	H	phosphorothioate
49	C41	OH	phosphorodithioate
50	C42	OMe	phosphorodithioate
51	C43	F	phosphorodithioate
52	C44	Cl	phosphorodithioate
53	C45	Br	phosphorodithioate
54	C46	I	phosphorodithioate
55	C47	NH ₂	phosphorodithioate
56	C48	H	phosphorodithioate
57	C49	OH	methylphosphonate
58	C50	OMe	methylphosphonate
59	C51	F	methylphosphonate
60	C52	Cl	methylphosphonate
61	C53	Br	methylphosphonate
62	C54	I	methylphosphonate
63	C55	NH ₂	methylphosphonate
64	C56	H	methylphosphonate
65	C57	OH	boranophosphonate
66	C58	OMe	boranophosphonate
67	C59	F	boranophosphonate
68	C60	Cl	boranophosphonate
69	C61	Br	boranophosphonate
70	C62	I	boranophosphonate
71	C63	NH ₂	boranophosphonate
72	C64	H	boranophosphonate
73	D1	OH	OH
74	D2	OMe	OH
75	D3	F	OH
76	D4	Cl	OH
77	D5	Br	OH
78	D6	I	OH

guanosine

guanosine

guanosine

guanosine

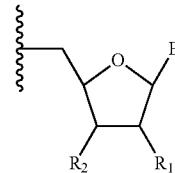
guanosine

US 10,337,001 B2

15

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.

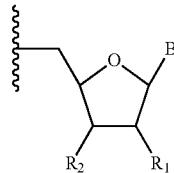


#	R ₁	R ₂	B
D7	NH ₂	OH	guanosine
D8	H	OH	guanosine
D9	OH	phosphodiester	guanosine
D10	OMe	phosphodiester	guanosine
D11	F	phosphodiester	guanosine
D12	Cl	phosphodiester	guanosine
D13	Br	phosphodiester	guanosine
D14	I	phosphodiester	guanosine
D15	NH ₂	phosphodiester	guanosine
D16	H	phosphodiester	guanosine
D17	OH	phosphonoacetate	guanosine
D18	OMe	phosphonoacetate	guanosine
D19	F	phosphonoacetate	guanosine
D20	Cl	phosphonoacetate	guanosine
D21	Br	phosphonoacetate	guanosine
D22	I	phosphonoacetate	guanosine
D23	NH ₂	phosphonoacetate	guanosine
D24	H	phosphonoacetate	guanosine
D25	OH	thiophosphonoacetate	guanosine
D26	OMe	thiophosphonoacetate	guanosine
D27	F	thiophosphonoacetate	guanosine
D28	Cl	thiophosphonoacetate	guanosine
D29	Br	thiophosphonoacetate	guanosine
D30	I	thiophosphonoacetate	guanosine
D31	NH ₂	thiophosphonoacetate	guanosine
D32	H	thiophosphonoacetate	guanosine
D33	OH	phosphorothioate	guanosine
D34	OMe	phosphorothioate	guanosine
D35	F	phosphorothioate	guanosine
D36	Cl	phosphorothioate	guanosine
D37	Br	phosphorothioate	guanosine
D38	I	phosphorothioate	guanosine
D39	NH ₂	phosphorothioate	guanosine
D40	H	phosphorothioate	guanosine
D41	OH	phosphorodithioate	guanosine
D42	OMe	phosphorodithioate	guanosine
D43	F	phosphorodithioate	guanosine
D44	Cl	phosphorodithioate	guanosine
D45	Br	phosphorodithioate	guanosine
D46	I	phosphorodithioate	guanosine
D47	NH ₂	phosphorodithioate	guanosine
D48	H	phosphorodithioate	guanosine
D49	OH	methylphosphonate	guanosine
D50	OMe	methylphosphonate	guanosine
D51	F	methylphosphonate	guanosine
D52	Cl	methylphosphonate	guanosine
D53	Br	methylphosphonate	guanosine
D54	I	methylphosphonate	guanosine
D55	NH ₂	methylphosphonate	guanosine
D56	H	methylphosphonate	guanosine
D57	OH	boranophosphonate	guanosine
D58	OMe	boranophosphonate	guanosine
D59	F	boranophosphonate	guanosine
D60	Cl	boranophosphonate	guanosine
D61	Br	boranophosphonate	guanosine
D62	I	boranophosphonate	guanosine
D63	NH ₂	boranophosphonate	guanosine
D64	H	boranophosphonate	guanosine
E1	OH	2-thiouridine	2-thiouridine
E2	OMe	OH	2-thiouridine
E3	F	OH	2-thiouridine
E4	Cl	OH	2-thiouridine
E5	Br	OH	2-thiouridine
E6	I	OH	2-thiouridine
E7	NH ₂	OH	2-thiouridine
E8	H	OH	2-thiouridine

16

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.



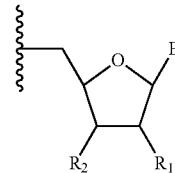
5			
10			
	#	R ₁	R ₂
			B
15	E9	OH	phosphodiester
	E10	OMe	phosphodiester
	E11	F	phosphodiester
	E12	Cl	phosphodiester
	E13	Br	phosphodiester
	E14	I	phosphodiester
	E15	NH ₂	phosphodiester
	E16	H	phosphodiester
20	E17	OH	phosphonoacetate
	E18	OMe	phosphonoacetate
	E19	F	phosphonoacetate
	E20	Cl	phosphonoacetate
	E21	Br	phosphonoacetate
	E22	I	phosphonoacetate
25	E23	NH ₂	phosphonoacetate
	E24	H	phosphonoacetate
	E25	OH	thiophosphonoacetate
	E26	OMe	thiophosphonoacetate
	E27	F	thiophosphonoacetate
	E28	Cl	thiophosphonoacetate
	E29	Br	thiophosphonoacetate
30	E30	I	thiophosphonoacetate
	E31	NH ₂	thiophosphonoacetate
	E32	H	thiophosphonoacetate
	E33	OH	phosphorothioate
	E34	OMe	phosphorothioate
	E35	F	phosphorothioate
35	E36	Cl	phosphorothioate
	E37	Br	phosphorothioate
	E38	I	phosphorothioate
	E39	NH ₂	phosphorothioate
	E40	H	phosphorothioate
	E41	OH	phosphorodithioate
40	E42	OMe	phosphorodithioate
	E43	F	phosphorodithioate
	E44	Cl	phosphorodithioate
	E45	Br	phosphorodithioate
	E46	I	phosphorodithioate
	E47	NH ₂	phosphorodithioate
	E48	H	phosphorodithioate
45	E49	OH	methylphosphonate
	E50	OMe	methylphosphonate
	E51	F	methylphosphonate
	E52	Cl	methylphosphonate
	E53	Br	methylphosphonate
	E54	I	methylphosphonate
50	E55	NH ₂	methylphosphonate
	E56	H	methylphosphonate
	E57	OH	boranophosphonate
	E58	OMe	boranophosphonate
	E59	F	boranophosphonate
	E60	Cl	boranophosphonate
55	E61	Br	boranophosphonate
	E62	I	boranophosphonate
	E63	NH ₂	boranophosphonate
	E64	H	boranophosphonate
	F1	OH	OH
	F2	OMe	OH
60	F3	F	OH
	F4	Cl	OH
	F5	Br	OH
	F6	I	OH
	F7	NH ₂	OH
65	F8	H	OH
	F9	OH	phosphodiester
	F10	OMe	phosphodiester

US 10,337,001 B2

17

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.

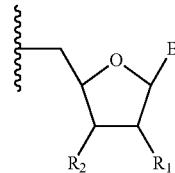


#	R ₁	R ₂	B
F11	F	phosphodiester	4-thiouridine
F12	Cl	phosphodiester	4-thiouridine
F13	Br	phosphodiester	4-thiouridine
F14	I	phosphodiester	4-thiouridine
F15	NH ₂	phosphodiester	4-thiouridine
F16	H	phosphodiester	4-thiouridine
F17	OH	phosphonoacetate	4-thiouridine
F18	OMe	phosphonoacetate	4-thiouridine
F19	F	phosphonoacetate	4-thiouridine
F20	Cl	phosphonoacetate	4-thiouridine
F21	Br	phosphonoacetate	4-thiouridine
F22	I	phosphonoacetate	4-thiouridine
F23	NH ₂	phosphonoacetate	4-thiouridine
F24	H	phosphonoacetate	4-thiouridine
F25	OH	thiophosphonoacetate	4-thiouridine
F26	OMe	thiophosphonoacetate	4-thiouridine
F27	F	thiophosphonoacetate	4-thiouridine
F28	Cl	thiophosphonoacetate	4-thiouridine
F29	Br	thiophosphonoacetate	4-thiouridine
F30	I	thiophosphonoacetate	4-thiouridine
F31	NH ₂	thiophosphonoacetate	4-thiouridine
F32	H	thiophosphonoacetate	4-thiouridine
F33	OH	phosphorothioate	4-thiouridine
F34	OMe	phosphorothioate	4-thiouridine
F35	F	phosphorothioate	4-thiouridine
F36	Cl	phosphorothioate	4-thiouridine
F37	Br	phosphorothioate	4-thiouridine
F38	I	phosphorothioate	4-thiouridine
F39	NH ₂	phosphorothioate	4-thiouridine
F40	H	phosphorothioate	4-thiouridine
F41	OH	phosphorodithioate	4-thiouridine
F42	OMe	phosphorodithioate	4-thiouridine
F43	F	phosphorodithioate	4-thiouridine
F44	Cl	phosphorodithioate	4-thiouridine
F45	Br	phosphorodithioate	4-thiouridine
F46	I	phosphorodithioate	4-thiouridine
F47	NH ₂	phosphorodithioate	4-thiouridine
F48	H	phosphorodithioate	4-thiouridine
F49	OH	methylphosphonate	4-thiouridine
F50	OMe	methylphosphonate	4-thiouridine
F51	F	methylphosphonate	4-thiouridine
F52	Cl	methylphosphonate	4-thiouridine
F53	Br	methylphosphonate	4-thiouridine
F54	I	methylphosphonate	4-thiouridine
F55	NH ₂	methylphosphonate	4-thiouridine
F56	H	methylphosphonate	4-thiouridine
F57	OH	boranophosphonate	4-thiouridine
F58	OMe	boranophosphonate	4-thiouridine
F59	F	boranophosphonate	4-thiouridine
F60	Cl	boranophosphonate	4-thiouridine
F61	Br	boranophosphonate	4-thiouridine
F62	I	boranophosphonate	4-thiouridine
F63	NH ₂	boranophosphonate	4-thiouridine
F64	H	boranophosphonate	4-thiouridine
G1	OH	OH	2-aminoadenosine
G2	OMe	OH	2-aminoadenosine
G3	F	OH	2-aminoadenosine
G4	Cl	OH	2-aminoadenosine
G5	Br	OH	2-aminoadenosine
G6	I	OH	2-aminoadenosine
G7	NH ₂	OH	2-aminoadenosine
G8	H	OH	2-aminoadenosine
G9	OH	phosphodiester	2-aminoadenosine
G10	OMe	phosphodiester	2-aminoadenosine
G11	F	phosphodiester	2-aminoadenosine
G12	Cl	phosphodiester	2-aminoadenosine

18

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.



5			
10			
	#	R ₁	R ₂
			B
15	G13	Br	phosphodiester
	G14	I	phosphodiester
	G15	NH ₂	phosphodiester
	G16	H	phosphodiester
	G17	OH	phosphonoacetate
	G18	OMe	phosphonoacetate
	G19	F	phosphonoacetate
	G20	Cl	phosphonoacetate
20	G21	Br	phosphonoacetate
	G22	I	phosphonoacetate
	G23	NH ₂	phosphonoacetate
	G24	H	phosphonoacetate
	G25	OH	thiophosphonoacetate
	G26	OMe	thiophosphonoacetate
	G27	F	thiophosphonoacetate
	G28	Cl	thiophosphonoacetate
	G29	Br	thiophosphonoacetate
25	G30	I	thiophosphonoacetate
	G31	NH ₂	thiophosphonoacetate
	G32	H	thiophosphonoacetate
	G33	OH	phosphorothioate
30	G34	OMe	phosphorothioate
	G35	F	phosphorothioate
	G36	Cl	phosphorothioate
	G37	Br	phosphorothioate
	G38	I	phosphorothioate
	G39	NH ₂	phosphorothioate
35	G40	H	phosphorothioate
	G41	OH	phosphorodithioate
	G42	OMe	phosphorodithioate
	G43	F	phosphorodithioate
	G44	Cl	phosphorodithioate
	G45	Br	phosphorodithioate
40	G46	I	phosphorodithioate
	G47	NH ₂	phosphorodithioate
	G48	H	phosphorodithioate
	G49	OH	methylphosphonate
	G50	OMe	methylphosphonate
	G51	F	methylphosphonate
	G52	Cl	methylphosphonate
	G53	Br	methylphosphonate
	G54	I	methylphosphonate
	G55	NH ₂	methylphosphonate
45	G56	H	methylphosphonate
	G57	OH	boranophosphonate
	G58	OMe	boranophosphonate
	G59	F	boranophosphonate
	G60	Cl	boranophosphonate
	G61	Br	boranophosphonate
	G62	I	boranophosphonate
	G63	NH ₂	boranophosphonate
	G64	H	boranophosphonate
55	H1	OH	OH
	H2	OMe	OH
		F	OH
		Cl	OH
	H5	Br	OH
	H6	I	OH
60	H7	NH ₂	OH
	H8	H	OH
	H9	OH	phosphodiester
	H10	OMe	phosphodiester
	H11	F	phosphodiester
	H12	Cl	phosphodiester
65	H13	Br	phosphodiester
	H14	I	phosphodiester

US 10,337,001 B2

19

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.

#	R ₁	R ₂	B
H15	NH ₂	phosphodiester	7-deazaguanosine
H16	H	phosphodiester	7-deazaguanosine
H17	OH	phosphonoacetate	7-deazaguanosine
H18	OMe	phosphonoacetate	7-deazaguanosine
H19	F	phosphonoacetate	7-deazaguanosine
H20	Cl	phosphonoacetate	7-deazaguanosine
H21	Br	phosphonoacetate	7-deazaguanosine
H22	I	phosphonoacetate	7-deazaguanosine
H23	NH ₂	phosphonoacetate	7-deazaguanosine
H24	H	phosphonoacetate	7-deazaguanosine
H25	OH	thiophosphonoacetate	7-deazaguanosine
H26	OMe	thiophosphonoacetate	7-deazaguanosine
H27	F	thiophosphonoacetate	7-deazaguanosine
H28	Cl	thiophosphonoacetate	7-deazaguanosine
H29	Br	thiophosphonoacetate	7-deazaguanosine
H30	I	thiophosphonoacetate	7-deazaguanosine
H31	NH ₂	thiophosphonoacetate	7-deazaguanosine
H32	H	thiophosphonoacetate	7-deazaguanosine
H33	OH	phosphorothioate	7-deazaguanosine
H34	OMe	phosphorothioate	7-deazaguanosine
H35	F	phosphorothioate	7-deazaguanosine
H36	Cl	phosphorothioate	7-deazaguanosine
H37	Br	phosphorothioate	7-deazaguanosine
H38	I	phosphorothioate	7-deazaguanosine
H39	NH ₂	phosphorothioate	7-deazaguanosine
H40	H	phosphorothioate	7-deazaguanosine
H41	OH	phosphorodithioate	7-deazaguanosine
H42	OMe	phosphorodithioate	7-deazaguanosine
H43	F	phosphorodithioate	7-deazaguanosine
H44	Cl	phosphorodithioate	7-deazaguanosine
H45	Br	phosphorodithioate	7-deazaguanosine
H46	I	phosphorodithioate	7-deazaguanosine
H47	NH ₂	phosphorodithioate	7-deazaguanosine
H48	H	phosphorodithioate	7-deazaguanosine
H49	OH	methylphosphonate	7-deazaguanosine
H50	OMe	methylphosphonate	7-deazaguanosine
H51	F	methylphosphonate	7-deazaguanosine
H52	Cl	methylphosphonate	7-deazaguanosine
H53	Br	methylphosphonate	7-deazaguanosine
H54	I	methylphosphonate	7-deazaguanosine
H55	NH ₂	methylphosphonate	7-deazaguanosine
H56	H	methylphosphonate	7-deazaguanosine
H57	OH	boranophosphonate	7-deazaguanosine
H58	OMe	boranophosphonate	7-deazaguanosine
H59	F	boranophosphonate	7-deazaguanosine
H60	Cl	boranophosphonate	7-deazaguanosine
H61	Br	boranophosphonate	7-deazaguanosine
H62	I	boranophosphonate	7-deazaguanosine
H63	NH ₂	boranophosphonate	7-deazaguanosine
H64	H	boranophosphonate	7-deazaguanosine
I1	OH	OH	inosine
I2	OMe	OH	inosine
I3	F	OH	inosine
I4	Cl	OH	inosine
I5	Br	OH	inosine
I6	I	OH	inosine
I7	NH ₂	OH	inosine
I8	H	OH	inosine
I9	OH	phosphodiester	inosine
I10	OMe	phosphodiester	inosine
I11	F	phosphodiester	inosine
I12	Cl	phosphodiester	inosine
I13	Br	phosphodiester	inosine
I14	I	phosphodiester	inosine
I15	NH ₂	phosphodiester	inosine
I16	H	phosphodiester	inosine

20

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.

5			
10			
#	R ₁	R ₂	B
I17	OH	phosphonoacetate	inosine
I18	OMe	phosphonoacetate	inosine
I19	F	phosphonoacetate	inosine
I20	Cl	phosphonoacetate	inosine
I21	Br	phosphonoacetate	inosine
I22	I	phosphonoacetate	inosine
I23	NH ₂	phosphonoacetate	inosine
I24	H	phosphonoacetate	inosine
I25	OH	thiophosphonoacetate	inosine
I26	OMe	thiophosphonoacetate	inosine
I27	F	thiophosphonoacetate	inosine
I28	Cl	thiophosphonoacetate	inosine
I29	Br	thiophosphonoacetate	inosine
I30	I	thiophosphonoacetate	inosine
I31	NH ₂	thiophosphonoacetate	inosine
I32	H	thiophosphonoacetate	inosine
I33	OH	phosphorothioate	inosine
I34	OMe	phosphorothioate	inosine
I35	F	phosphorothioate	inosine
I36	Cl	phosphorothioate	inosine
I37	Br	phosphorothioate	inosine
I38	I	phosphorothioate	inosine
I39	NH ₂	phosphorothioate	inosine
I40	H	phosphorothioate	inosine
I41	OH	phosphorodithioate	inosine
I42	OMe	phosphorodithioate	inosine
I43	F	phosphorodithioate	inosine
I44	Cl	phosphorodithioate	inosine
I45	Br	phosphorodithioate	inosine
I46	I	phosphorodithioate	inosine
I47	NH ₂	phosphorodithioate	inosine
I48	H	phosphorodithioate	inosine
I49	OH	methylphosphonate	inosine
I50	OMe	methylphosphonate	inosine
I51	F	methylphosphonate	inosine
I52	Cl	methylphosphonate	inosine
I53	Br	methylphosphonate	inosine
I54	I	methylphosphonate	inosine
I55	NH ₂	methylphosphonate	inosine
I56	H	methylphosphonate	inosine
I57	OH	boranophosphonate	inosine
I58	OMe	boranophosphonate	inosine
I59	F	boranophosphonate	inosine
I60	Cl	boranophosphonate	inosine
I61	Br	boranophosphonate	inosine
I62	I	boranophosphonate	inosine
I63	NH ₂	boranophosphonate	inosine
I64	H	boranophosphonate	inosine
J1	OH	OH	5-methylcytidine
J2	OMe	OH	5-methylcytidine
J3	F	OH	5-methylcytidine
J4	Cl	OH	5-methylcytidine
J5	Br	OH	5-methylcytidine
J6	I	OH	5-methylcytidine
J7	NH ₂	OH	5-methylcytidine
J8	H	OH	5-methylcytidine
J9	OH	phosphodiester	5-methylcytidine
J10	OMe	phosphodiester	5-methylcytidine
J11	F	phosphodiester	5-methylcytidine
J12	Cl	phosphodiester	5-methylcytidine
J13	Br	phosphodiester	5-methylcytidine
J14	I	phosphodiester	5-methylcytidine
J15	NH ₂	phosphodiester	5-methylcytidine
J16	H	phosphodiester	5-methylcytidine
J17	OH	phosphonoacetate	5-methylcytidine
J18	OMe	phosphonoacetate	5-methylcytidine

US 10,337,001 B2

21

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.

#	R ₁	R ₂	B
J19	F	phosphonoacetate	5-methylcytidine
J20	Cl	phosphonoacetate	5-methylcytidine
J21	Br	phosphonoacetate	5-methylcytidine
J22	I	phosphonoacetate	5-methylcytidine
J23	NH ₂	phosphonoacetate	5-methylcytidine
J24	H	phosphonoacetate	5-methylcytidine
J25	OH	thiophosphonoacetate	5-methylcytidine
J26	OMe	thiophosphonoacetate	5-methylcytidine
J27	F	thiophosphonoacetate	5-methylcytidine
J28	Cl	thiophosphonoacetate	5-methylcytidine
J29	Br	thiophosphonoacetate	5-methylcytidine
J30	I	thiophosphonoacetate	5-methylcytidine
J31	NH ₂	thiophosphonoacetate	5-methylcytidine
J32	H	thiophosphonoacetate	5-methylcytidine
J33	OH	phosphorothioate	5-methylcytidine
J34	OMe	phosphorothioate	5-methylcytidine
J35	F	phosphorothioate	5-methylcytidine
J36	Cl	phosphorothioate	5-methylcytidine
J37	Br	phosphorothioate	5-methylcytidine
J38	I	phosphorothioate	5-methylcytidine
J39	NH ₂	phosphorothioate	5-methylcytidine
J40	H	phosphorothioate	5-methylcytidine
J41	OH	phosphorodithioate	5-methylcytidine
J42	OMe	phosphorodithioate	5-methylcytidine
J43	F	phosphorodithioate	5-methylcytidine
J44	Cl	phosphorodithioate	5-methylcytidine
J45	Br	phosphorodithioate	5-methylcytidine
J46	I	phosphorodithioate	5-methylcytidine
J47	NH ₂	phosphorodithioate	5-methylcytidine
J48	H	phosphorodithioate	5-methylcytidine
J49	OH	methylphosphonate	5-methylcytidine
J50	OMe	methylphosphonate	5-methylcytidine
J51	F	methylphosphonate	5-methylcytidine
J52	Cl	methylphosphonate	5-methylcytidine
J53	Br	methylphosphonate	5-methylcytidine
J54	I	methylphosphonate	5-methylcytidine
J55	NH ₂	methylphosphonate	5-methylcytidine
J56	H	methylphosphonate	5-methylcytidine
J57	OH	boranophosphonate	5-methylcytidine
J58	OMe	boranophosphonate	5-methylcytidine
J59	F	boranophosphonate	5-methylcytidine
J60	Cl	boranophosphonate	5-methylcytidine
J61	Br	boranophosphonate	5-methylcytidine
J62	I	boranophosphonate	5-methylcytidine
J63	NH ₂	boranophosphonate	5-methylcytidine
J64	H	boranophosphonate	5-methylcytidine
K1	OH	OH	5-aminoallyluridine
K2	OMe	OH	5-aminoallyluridine
K3	F	OH	5-aminoallyluridine
K4	Cl	OH	5-aminoallyluridine
K5	Br	OH	5-aminoallyluridine
K6	I	OH	5-aminoallyluridine
K7	NH ₂	OH	5-aminoallyluridine
K8	H	OH	5-aminoallyluridine
K9	OH	phosphodiester	5-aminoallyluridine
K10	OMe	phosphodiester	5-aminoallyluridine
K11	F	phosphodiester	5-aminoallyluridine
K12	Cl	phosphodiester	5-aminoallyluridine
K13	Br	phosphodiester	5-aminoallyluridine
K14	I	phosphodiester	5-aminoallyluridine
K15	NH ₂	phosphodiester	5-aminoallyluridine
K16	H	phosphodiester	5-aminoallyluridine
K17	OH	phosphonoacetate	5-aminoallyluridine
K18	OMe	phosphonoacetate	5-aminoallyluridine
K19	F	phosphonoacetate	5-aminoallyluridine
K20	Cl	phosphonoacetate	5-aminoallyluridine

22

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.

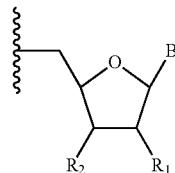
5			
10			
#	R ₁	R ₂	B
K21	Br	phosphonoacetate	5-aminoallyluridine
K22	I	phosphonoacetate	5-aminoallyluridine
15	K23	NH ₂	phosphonoacetate
K24	H	phosphonoacetate	5-aminoallyluridine
K25	OH	thiophosphonoacetate	5-aminoallyluridine
K26	OMe	thiophosphonoacetate	5-aminoallyluridine
K27	F	thiophosphonoacetate	5-aminoallyluridine
K28	Cl	thiophosphonoacetate	5-aminoallyluridine
20	K29	Br	thiophosphonoacetate
K30	I	thiophosphonoacetate	5-aminoallyluridine
K31	NH ₂	thiophosphonoacetate	5-aminoallyluridine
K32	H	thiophosphonoacetate	5-aminoallyluridine
K33	OH	phosphorothioate	5-aminoallyluridine
K34	OMe	phosphorothioate	5-aminoallyluridine
25	K35	F	phosphorothioate
K36	Cl	phosphorothioate	5-aminoallyluridine
K37	Br	phosphorothioate	5-aminoallyluridine
K38	I	phosphorothioate	5-aminoallyluridine
K39	NH ₂	phosphorothioate	5-aminoallyluridine
K40	H	phosphorothioate	5-aminoallyluridine
30	K41	OH	phosphorodithioate
K42	OMe	phosphorodithioate	5-aminoallyluridine
K43	F	phosphorodithioate	5-aminoallyluridine
K44	Cl	phosphorodithioate	5-aminoallyluridine
K45	Br	phosphorodithioate	5-aminoallyluridine
K46	I	phosphorodithioate	5-aminoallyluridine
K47	NH ₂	phosphorodithioate	5-aminoallyluridine
35	K48	H	phosphorodithioate
K49	OH	methylphosphonate	5-aminoallyluridine
K50	OMe	methylphosphonate	5-aminoallyluridine
K51	F	methylphosphonate	5-aminoallyluridine
K52	Cl	methylphosphonate	5-aminoallyluridine
K53	Br	methylphosphonate	5-aminoallyluridine
40	K54	I	methylphosphonate
K55	NH ₂	methylphosphonate	5-aminoallyluridine
K56	H	methylphosphonate	5-aminoallyluridine
K57	OH	boranophosphonate	5-aminoallyluridine
K58	OMe	boranophosphonate	5-aminoallyluridine
K59	F	boranophosphonate	5-aminoallyluridine
45	K60	Cl	boranophosphonate
K61	Br	boranophosphonate	5-aminoallyluridine
K62	I	boranophosphonate	5-aminoallyluridine
K63	NH ₂	boranophosphonate	5-aminoallyluridine
K64	H	boranophosphonate	5-aminoallyluridine
L1	OH	OH	5-methyluridine
L2	OMe	OH	5-methyluridine
50	L3	F	5-methyluridine
L4	Cl	OH	5-methyluridine
L5	Br	OH	5-methyluridine
L6	I	OH	5-methyluridine
L7	NH ₂	OH	5-methyluridine
L8	H	OH	5-methyluridine
55	L9	OH	phosphodiester
L10	OMe	phosphodiester	5-methyluridine
L11	F	phosphodiester	5-methyluridine
L12	Cl	phosphodiester	5-methyluridine
L13	Br	phosphodiester	5-methyluridine
L14	I	phosphodiester	5-methyluridine
60	L15	NH ₂	phosphodiester
L16	H	phosphodiester	5-methyluridine
L17	OH	phosphonoacetate	5-methyluridine
L18	OMe	phosphonoacetate	5-methyluridine
L19	F	phosphonoacetate	5-methyluridine
L20	Cl	phosphonoacetate	5-methyluridine
65	L21	Br	phosphonoacetate
L22	I	phosphonoacetate	5-methyluridine

US 10,337,001 B2

23

TABLE 2-continued

Exemplary modified nucleotides contained in a synthetic guide sequence.



#	R ₁	R ₂	B
L23	NH ₂	phosphonoacetate	5-methyluridine
L24	H	phosphonoacetate	5-methyluridine
L25	OH	thiophosphonoacetate	5-methyluridine
L26	OMe	thiophosphonoacetate	5-methyluridine
L27	F	thiophosphonoacetate	5-methyluridine
L28	Cl	thiophosphonoacetate	5-methyluridine
L29	Br	thiophosphonoacetate	5-methyluridine
L30	I	thiophosphonoacetate	5-methyluridine
L31	NH ₂	thiophosphonoacetate	5-methyluridine
L32	H	thiophosphonoacetate	5-methyluridine
L33	OH	phosphorothioate	5-methyluridine
L34	OMe	phosphorothioate	5-methyluridine
L35	F	phosphorothioate	5-methyluridine
L36	Cl	phosphorothioate	5-methyluridine
L37	Br	phosphorothioate	5-methyluridine
L38	I	phosphorothioate	5-methyluridine
L39	NH ₂	phosphorothioate	5-methyluridine
L40	H	phosphorothioate	5-methyluridine
L41	OH	phosphorodithioate	5-methyluridine
L42	OMe	phosphorodithioate	5-methyluridine
L43	F	phosphorodithioate	5-methyluridine
L44	Cl	phosphorodithioate	5-methyluridine
L45	Br	phosphorodithioate	5-methyluridine
L46	I	phosphorodithioate	5-methyluridine
L47	NH ₂	phosphorodithioate	5-methyluridine
L48	H	phosphorodithioate	5-methyluridine
L49	OH	methylphosphonate	5-methyluridine
L50	OMe	methylphosphonate	5-methyluridine
L51	F	methylphosphonate	5-methyluridine
L52	Cl	methylphosphonate	5-methyluridine
L53	Br	methylphosphonate	5-methyluridine
L54	I	methylphosphonate	5-methyluridine
L55	NH ₂	methylphosphonate	5-methyluridine
L56	H	methylphosphonate	5-methyluridine
L57	OH	boranophosphonate	5-methyluridine
L58	OMe	boranophosphonate	5-methyluridine
L59	F	boranophosphonate	5-methyluridine
L60	Cl	boranophosphonate	5-methyluridine
L61	Br	boranophosphonate	5-methyluridine
L62	I	boranophosphonate	5-methyluridine
L63	NH ₂	boranophosphonate	5-methyluridine
L64	H	boranophosphonate	5-methyluridine

R₁ = OH or 2' modificationR₂ = OH or internucleotide linkage

B = base

As described herein, certain unnatural base pairs (e.g., isoG and isoC, Z base and P base; see Zhang et al. (2015) *J. Am. Chem. Soc.*) may be advantageous for affecting the thermostability of the guide RNA secondary structure. These modifications can be used to prevent misfolding of the guide RNA scaffold with other domains of a guide RNA sequence.

Recent guide RNA:Cas9 protein structural information (FIG. 10, as reported in Jiang et al. 2015, *Science*,) and in vivo/in vitro functional mutation studies (see, e.g., Briner et al. 2014, *Mol. Cell*, 56, 333-9) indicate that the guide RNA scaffold is predominantly structurally conserved. This reinforces the importance of correct folding of the conserved domain of guide RNAs for functionality with Cas9. FIG. 10 shows the guide RNA scaffold secondary structure, displaying interactions with amino acids of Cas9. Most of the guide RNA nitrogenous bases are not involved in binding interactions with Cas9 protein.

24

The flanking sequences of the sgRNA scaffold increase the likelihood of misfolding and hence malfunction. The 20 nt guide targeting sequence, 5' of the scaffold region, is user-specified for each target, thus the likelihood of misfolding is variable or target-specific. Also, many emerging CRISPR-Cas applications append functional sequences 3' of the scaffold, such as CRISPRdisplay (Schechner et al., *Nat. Methods* 2015) and CRISPR-i/-a (Chen et al., *Cell* 2013), which are riboswitches or aptamers that also need to correctly and independently fold to function properly. To ensure that each of the functional domains (i.e. targeting guide, scaffold, aptamer) of a given sgRNA folds in a modular, independent manner, the structurally conserved scaffold base pairs can be substituted with unnatural, orthogonal base pairs (e.g., isoG and isoC; Z base and P base), and in some embodiments, substituted exclusively with unnatural, orthogonal base pairs, this ensures that the sgRNA scaffold sequences will not stably interact in a secondary structure with elements of the target-pairing guide sequence or other non-native domains incorporated in the guide RNA such as any aptamer sequences or any non-native 5' or 3' overhangs on the guide RNA. Alternatively, the unnatural, orthogonal base pairs mentioned above could be incorporated in any non-native overhangs or aptamers that may be present, thus to prevent secondary structures involving misfolding of the scaffold sequence(s).

B. Guide RNA with at Least One Modification

In one aspect, the present technology provides a guide RNA having at least one modification, constituting a modified gRNA.

In certain embodiments, the modified gRNA comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 modified nucleotides. In other embodiments, the modified gRNA comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified nucleotides. In certain embodiments, all nucleotides are modified. In certain embodiments, all the modifications are the same. In certain embodiments, all the modified nucleotides have the same type of modification. In certain embodiments, the modified gRNA comprises a combination of differently modified nucleotides. In certain embodiments, the modified gRNA comprises two or more modified nucleotides. In certain embodiments, the modified gRNA comprises three or more modified nucleotides. In certain embodiments, the modified nucleotides are arranged contiguously. In certain embodiments, the modified gRNA comprises at least one contiguous stretch of modified nucleotides. In certain embodiments, the modified gRNA comprises a contiguous stretch of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 modified nucleotides. Each modified nucleotide may independently comprise one or more types of modifications. In certain embodiments, no modified nucleotides are contiguous, or some but not all are contiguous, in the sequence of the modified gRNA.

In certain embodiments, the modification is within the 5' portion of the guide RNA. In certain embodiments, the modification is within the first five (5) nucleotides of the 5' portion of the guide RNA. In certain embodiments, the modification is within the first three (3) nucleotides of the 5' portion of the guide RNA. In certain embodiments, the modification is within the 3' portion of the guide RNA. In

US 10,337,001 B2

25

certain embodiments, the modification is within the last five (5) nucleotides of the 3' portion of the guide RNA. In certain embodiments, the modification is within the last three (3) nucleotides of the 3' portion of the guide RNA. In certain embodiments, the modification is within the internal region (i.e., between the 5' end and the 3' end) of the guide RNA.

In certain embodiments, the modification is incorporated in the 5' portion or the 3' portion of the guide RNA, particularly within the first 5 or 10 nucleotides of the 5' portion or within the last 5 or 10 nucleotides of the 3' portion to, for example, protect the RNA from degradation by nucleases or for other purposes. In some other embodiments, the modification is in both the 5' portion and the 3' portion of the guide RNA, particularly within the first 5 or 10 nucleotides of the 5' portion and within the last 5 or 10 nucleotides of the 3' portion to, for example, protect the RNA from degradation by nucleases or for other purposes. In certain embodiments, more than one type of modification is present in both the 5' portion and the 3' portion of the guide RNA. In certain embodiments, the modifications are located at the 5' end, at the 3' end, and within the internal sequence of the guide RNA. In certain embodiments, a guide RNA comprises 40 or fewer, alternatively 20 or fewer, alternatively 15 or fewer, alternatively 10 or fewer, alternatively 5 or fewer, alternatively 3 or fewer deoxyribonucleotide residues in the 5' or 3' portion of the guide RNA.

In certain embodiments, the modification is within the crRNA segment of the guide RNA. In certain embodiments, the modification is within the guide sequence of the crRNA. In certain embodiments, the modification is within the first five (5) nucleotides of the crRNA segment. In certain embodiments, the modification is within the first three (3) nucleotides of the crRNA segment. In certain embodiments, the modification is within a 5'-overhang on the crRNA segment. In certain embodiments, the modification is within the tracrRNA segment of the guide RNA. In certain embodiments, the modification is within the last five (5) nucleotides of the tracrRNA segment of the guide RNA. In certain embodiments, the modification is within the last three (3) nucleotides of the tracrRNA segment of the guide RNA. In certain embodiments, when the guide RNA is a single guide RNA, the modification is located within the loop of the guide RNA. In certain embodiments, one or more modifications is within the loop L region. In certain embodiments, the modification comprises a dye, a non-fluorescent label, or a tag conjugated to a linker incorporated between two nucleotides as described above, for example by conjugation to a 2-(3-(dye/label/tag-amido)propanamido)propane-1,3-diol bis(phosphodiester) linker or to a modified base of a nucleotide m in the loop or L region.

In certain embodiments, the modification comprises an end modification, such as a 5' end modification or a 3' end modification. Examples of end modifications include, but are not limited to phosphorylation (as natural phosphate or polyphosphate or as modified phosphonate groups such as for example, alkylphosphonate, phosphonocarboxylate, phosphonoacetate, boranophosphonate, phosphorothioate, phosphorodithioate and the like), biotinylation, conjugating or conjugated molecules, linkers, dyes, labels, tags, functional groups (such as for example but not limited to 5'-amino, 5'-thio, 5'-amido, 5'-carboxy and the like), inverted linkages, or hydrocarbon moieties which may comprise ether, polyethylene glycol (PEG), ester, hydroxyl, aryl, halo, phosphodiester, bicyclic, heterocyclic or other organic functional group. In certain embodiments, the end modification comprises dimethoxytrityl.

26

In certain embodiments, the modification comprises a modified base. As used herein, "unmodified" bases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). Examples of modified bases include, but are not limited to, synthetic and natural bases such as 2-thioU, 2-thioG, 4-thioU, 6-thioG, 2-aminoA, 2-aminop, pseudouracil, hypoxanthine, 7-deazaguanine, 7-deaza-8-azaguanine, 7-deazaadenine, 7-deaza-8-azaadenine, 5-methylC, 5-methylU, 5-hydroxymethylcytosine, 5-hydroxymethyluracil, 5,6-dehydrouracil, 5-propynylcytosine, 5-propynyluracil, 5-ethynylcytosine, 5-ethynyluracil, 5-allylU, 5-allylC, 5-aminoallyl-uracil, and 5-aminoallyl-cytosine. In certain embodiments, the modification comprises an abasic nucleotide. In certain embodiments, the modification comprises a nonstandard purine or pyrimidine structure, such as Z or P, isoC or isoG, UNA, 5-methylpyrimidine, x(A,G,C,T) or y(A,G,C,T). In certain embodiments, the modified gRNA comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 modified bases. In other embodiments, the modified gRNA comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified bases. In certain embodiments, all bases in a gRNA are modified.

In certain embodiments, the modification comprises a modified sugar. Examples of modified sugars include, but are not limited to, sugars having modifications, at the 21 position or modifications at the 4' position. For example, in certain embodiments, the sugar comprises 2'-O—C₁₋₄alkyl, such as 2'-O-methyl (2'-OMe). In certain embodiments, the sugar comprises 2'-O—C₁₋₃alkyl-O—C₁₋₃alkyl, such as 2'-methoxyethoxy (2'-O—CH₂CH₂OCH₃) also known as 2'-O-(2-methoxyethyl) or 2'-MOE. In certain embodiments, the sugar comprises 2'-halo, such as 2'-F, 2'-Br, 2'-Cl, or 2'-I. In certain embodiments, the sugar comprises 2'-NH₂. In certain embodiments, the sugar comprises 2'-H (e.g., a deoxynucleotide). In certain embodiments, the sugar comprises 2'-arabino or 2'-F-arabino. In certain embodiments, the sugar comprises 2'-LNA or 2'-ULNA. In certain embodiments, the sugar comprises a 4'-thioribosyl. In certain embodiments, the modified gRNA comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 modified sugars. In other embodiments, the modified gRNA comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified sugars. In certain embodiments, all sugars in a gRNA are modified.

In certain embodiments, the modification comprises a modified backbone (i.e., an internucleotide linkage other than a natural phosphodiester). Examples of modified internucleotide linkages include, but are not limited to a phosphorothioate internucleotide linkage, a chiral phosphorothioate internucleotide linkage, a phosphorodithioate internucleotide linkage, a boranophosphonate internucleotide linkage, a C₁₋₄alkyl phosphonate internucleotide linkage such as a methylphosphonate internucleotide linkage, a boranophosphonate internucleotide linkage, a phosphonocarboxylate internucleotide linkage such as a phosphonoacetate internucleotide linkage, a phosphonocarboxylate ester internucleotide linkage such as a phosphonoacetate ester internucleotide linkage, a thiophosphonocarboxylate internucleotide linkage such as for example a thiophosphonoacetate internucleotide linkage, a thiophosphonocarboxylate ester internucleotide linkage such as a thiophosphonoacetate ester internucleotide linkage. Various salts, mixed salts and

US 10,337,001 B2

27

free acid forms are also included. In certain embodiments, the modified gRNA comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 modified internucleotide linkages. In other embodiments, the modified gRNA comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified internucleotide linkages. In certain embodiments, all internucleotide linkages in a gRNA are modified.

In certain embodiments, the modification is a 2'-O—C₁-alkyl, 2'-H, 2'-O—C₁₋₃alkyl-O—C₁₋₃alkyl, 2'-F, 2'-NH₂, 2'-arabino, 2'-F-arabino, 2'-LNA, 2'-ULNA, 4'-thioribosyl, 2-thioU, 2-thioC, 4-thioU, 6-thioG, 2-aminoA, 2-aminoP, pseudouracil, hypoxanthine, 7-deazaguanine, 7-deaza-8-azaguanine, 7-deazaadenine, 7-deaza-8-azaadenine, 5-MeC, 5-MeU, 5-hydroxymethylcytosine, 5-hydroxymethyluracil, 5,6-dehydrouracil, 5-propynylcytosine, 5-propynyluracil, 5-ethynylcytosine, 5-ethynyluracil, 5-allylU, 5-allylC, 5-aminoallyl-uracil, 5-aminoallyl-cytosine, an abasic nucleotide, Z, P, UNA, isoC, isoG, 5-methyl-pyrimidine, x(A,G,C,T), y(A,G,C,T), a 3'-phosphorothioate group, a 3'-phosphonoacetate group, a 3'-phosphonoacetate ester-group, a 3'-thiophosphonoacetate group, a 3'-thiophosphonoacetate ester group, a 3'-methylphosphonate group, a 3'-boranophosphonate group, a 3'-phosphorodithioate group, or combinations thereof.

In certain embodiments, the modified nucleotide comprises a 2'-O-methyl-3'-phosphorothioate. In certain embodiments, the modified nucleotide comprises a 2'-O-methyl-3'-phosphonoacetate. In certain embodiments, the modified nucleotide comprises a 2'-O-methyl-3'-thiophosphonoacetate. In certain embodiments, the modified nucleotide comprises a Z base. In certain embodiments, the modified nucleotide comprises a 2'-halo-3'-phosphorothioate. In certain embodiments, the modified nucleotide comprises a 2'-halo-3'-phosphonoacetate. In certain embodiments, the modified nucleotide comprises a 2'-halo-3'-thiophosphonoacetate. In certain embodiments, the modified nucleotide comprises a 2'-fluoro-3'-phosphorothioate. In certain embodiments, the modified nucleotide comprises a 2'-fluoro-3'-phosphonoacetate. In certain embodiments, the modified nucleotide comprises a 2'-fluoro-3'-thiophosphonoacetate.

In certain embodiments, the guide RNA comprises an oligonucleotide represented by Formula (I):

W-Y or Y-W (I)

wherein W represents a nucleotide or a stretch of nucleotides of the oligonucleotide comprising at least one modification and Y represents an unmodified portion of the oligonucleotide.

In certain embodiments, W is within the 5' portion of the guide RNA. In certain embodiments, W is at least partially within the first five (5) nucleotides of the 5' portion of the guide RNA. In certain embodiments, W is at least partially within the first three (3) nucleotides of the 5' portion of the guide RNA. In certain embodiments, W is within the 3' portion of the guide RNA. In certain embodiments, W is at least partially within the last five (5) nucleotides of the 3' portion of the guide RNA. In certain embodiments, W is at least partially within the last three (3) nucleotides of the 3' portion of the guide RNA. In certain embodiments, W is within the internal region (i.e., between the 5' end and the 3' end) of the guide RNA.

In certain embodiments, W comprises an end modification, such as a 5' end modification or a 3' end modification

28

as described above. In certain embodiments, the end modification comprises dimethoxytrityl.

In certain embodiments, W comprises a modified base as described above. In certain embodiments, W comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 modified bases. In other embodiments, W comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified bases. In certain embodiments, all bases in a gRNA are modified.

In certain embodiments, W comprises a modified sugar as described above. In certain embodiments, W comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 modified sugars. In other embodiments, W comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified sugars. In certain embodiments, all sugars in a gRNA are modified.

In certain embodiments, W comprises a modified backbone (i.e., an internucleotide linkage other than a phosphodiester) as described above. In certain embodiments, W comprises more than one, such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 modified internucleotide linkages. In other embodiments, W comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified internucleotide linkages. In certain embodiments, all internucleotide linkages in a gRNA are modified.

In certain embodiments, W comprises a 2'-O—C₁₋₄alkyl, 2'-H, 2'-O—C₁₋₃alkyl-O—C₁₋₃alkyl, 2'-F, 2'-NH₂, 2'-arabino, 2'-F-arabino, 2'-LNA, 2'-ULNA, 4'-thioribosyl, 2-thioU, 2-thioC, 4-thioU, 6-thioG, 2-aminoA, 2-aminoP, pseudouracil, hypoxanthine, 7-deazaguanine, 7-deaza-8-azaguanine, 7-deazaadenine, 7-deaza-8-azaadenine, 5-MeC, 5-MeU, 5-hydroxymethylcytosine, 5-hydroxymethyluracil, 5,6-dehydrouracil, 5-propynylcytosine, 5-propynyluracil, 5-ethynylcytosine, 5-ethynyluracil, 5-allylU, 5-allylC, 5-aminoallyl-uracil, 5-aminoallyl-cytosine, an abasic nucleotide, Z, P, UNA, isoC, isoG, 5-methyl-pyrimidine, x(A,G,C,T), y(A,G,C,T), a phosphorothioate internucleotide linkage, a phosphonoacetate internucleotide linkage, a thiophosphonoacetate internucleotide linkage, a thiophosphonate ester internucleotide linkage, a boranophosphonate internucleotide linkage, a phosphorodithioate internucleotide linkage, or combinations thereof.

In certain embodiments, W comprises a 2'-O-methyl and a 3'-phosphorothioate group on the same nucleotide. In certain embodiments, W comprises a 2'-O-methyl and a 3'-phosphonoacetate group on the same nucleotide. In certain embodiments, W comprises a 2'-O-methyl and 3'-thiophosphonoacetate group on the same nucleotide. In certain embodiments, W comprises a 2'-F and a 3'-phosphorothioate group on the same nucleotide. In certain embodiments, W comprises 2'-F and a 3'-phosphonoacetate group on the same nucleotide. In certain embodiments, W comprises a 2'-F and 3'-thiophosphonoacetate group on the same nucleotide.

In certain embodiments, W comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 modified nucleotides. In certain embodiments, each of the modified nucleotides comprises the same modification. In certain embodiments, W comprises a combination of variously modified nucleotides. In certain embodiments, W comprises

US 10,337,001 B2

29

two or more modified nucleotides. In certain embodiments, W comprises three or more modified nucleotides. In certain embodiments, the modified nucleotides are not arranged contiguously in the sequence, or at least not entirely, as one or more unmodified nucleotides may intercede. In certain embodiments, the modified nucleotides are arranged contiguously. In certain embodiments, W comprises at least one contiguous stretch of modified nucleotides. In certain embodiments, W comprises a contiguous stretch of at least three (3) modified nucleotides. In certain embodiments, W comprises a contiguous stretch of at least four (4) modified nucleotides. In certain embodiments, W comprises a contiguous stretch of at least five (5) modified nucleotides.

In certain embodiments, the guide RNA comprises an oligonucleotide represented by Formula (II):



wherein each N independently represents an unmodified ribonucleotide;

wherein each M represents a modified nucleotide and is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 3'-P(S) ribonucleotide, a 3'-PACE ribonucleotide, a 3'-thioPACE ribonucleotide, a 2'-O-methyl-3'-P(S)-ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, a Z nucleotide, and a 2'-deoxynucleotide;

wherein each M is at any position of the sequence of the guide RNA;

wherein any given M is the same or different from any other M, and any given N is the same or different from any other N; and

wherein each of m and n are independently selected from an integer between 0 and 219, provided that $50 < m+n \leq 220$, and m is not 0.

In some embodiments, $m+n < 150$.

In certain embodiments, each M is modified with one or more moieties independently selected from the group consisting of 2'-F, 2-thiouracil, 4-thiouracil, 2-aminoadenine, hypoxanthine, 5-methylcytosine, 5-methyluracil, 5-allylaminoouracil, squarate linkage, a triazolo linkage, and a 2-(4-butylamidofluorescein)propane-1,3-diol bis(phosphodiester) linkage. In some embodiments, M comprises a dye attached through a linker.

In certain embodiments, each M is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, and a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M is independently selected from the group consisting of a 2'-O-methyl-3'-PACE ribonucleotide and a 2'-O-methyl-3'-thioPACE ribonucleotide.

In certain embodiments, where $m > 1$, any given M is the same or different from any other M. In certain embodiments, where $m > 1$, each M has the same modification.

In certain embodiments, each M is a 2'-O-methyl-3'-PACE ribonucleotide, m is selected from an integer between 1 and 10, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a 2'-O-methyl-3'-PACE ribonucleotide, m is selected from an integer between 1 and 5, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a 2'-O-methyl-3'-PACE ribonucleotide, m is selected from an integer between 2 and 5, each N is independently selected from the group consisting of A,

30

U, C, and G, and n is selected from an integer between 1 and 148, provided $50 < m+n \leq 150$. In certain embodiments, m is 1. In certain embodiments, m is 2. In certain embodiments, m is 3. In certain embodiments, m is 4. In certain embodiments, m is 5.

In certain embodiments, each M is a 2'-O-methyl-3'-thioPACE ribonucleotide, m is selected from an integer between 1 and 10, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a 2'-O-methyl-3'-thioPACE ribonucleotide, m is selected from an integer between 1 and 5, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a 2'-O-methyl-3'-thioPACE ribonucleotide, m is selected from an integer between 2 and 5, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a 2'-O-methyl-3'-thioPACE ribonucleotide, m is selected from an integer between 1 and 148, provided $50 < m+n \leq 150$. In certain embodiments, m is 1. In certain embodiments, m is 2. In certain embodiments, m is 3. In certain embodiments, m is 4. In certain embodiments, m is 5.

In certain embodiments, each M is a 2'-O-methyl ribonucleotide, m is selected from an integer between 1 and 40, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a 2'-O-methyl ribonucleotide, m is selected from an integer between 1 and 25, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a 2'-O-methyl ribonucleotide, m is selected from an integer between 1 and 20, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, m is 1. In certain embodiments, m is 2. In certain embodiments, m is 3. In certain embodiments, m is 4. In certain embodiments, m is 5. In certain embodiments, m is 10. In certain embodiments, m is 15. In certain embodiments, m is 20. In certain embodiments, m is 30. In certain embodiments, m is 40.

In certain embodiments, each M is a 2'-deoxynucleotide, m is selected from an integer between 1 and 30, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is 2'-deoxynucleotide, m is selected from an integer between 1 and 20, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, m is 5. In certain embodiments, m is 10. In certain embodiments, m is 15. In certain embodiments, m is 20. In certain embodiments, m is 30.

In certain embodiments, each M is a 2'-O-methyl-3'-P(S) ribonucleotide, m is selected from an integer between 1 and 10, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a 2'-O-methyl-3'-P(S) ribonucleotide, m is selected from an integer between 1 and 5, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, m is 1.

US 10,337,001 B2

31

In certain embodiments, m is 2. In certain embodiments, m is 3. In certain embodiments, m is 4. In certain embodiments, m is 5.

In certain embodiments, each M is a Z nucleotide, m is selected from an integer between 1 and 10, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, each M is a Z nucleotide, m is selected from an integer between 1 and 5, each N is independently selected from the group consisting of A, U, C, and G, and n is selected from an integer between 1 and 149, provided $50 < m+n \leq 150$. In certain embodiments, m is 1. In certain embodiments, m is 2. In certain embodiments, m is 3. In certain embodiments, m is 4. In certain embodiments, m is 5.

In certain embodiments, the modification is a stability-altering modification. In certain embodiments, the modification increases nuclease resistance of the guide RNA relative to a guide RNA without the modification, thus it enhances the guide RNA stability. In certain embodiments, the stability-altering modification is a stability-enhancing modification. For example, in certain embodiments, the stability-enhancing modification comprises, a 2'-O-methyl or a 2'-O— C_{1-4} alkyl nucleotide. In certain embodiments, the stability-enhancing modification comprises a 2'-halo nucleotide, such as 2'-F, 2'-Br, 2'-Cl, or 2'-I. In certain embodiments, the stability-enhancing modification comprises a 2'MOE or a 2'-O— C_{1-3} alkyl-O— C_{1-3} alkyl. In certain embodiments, the stability-enhancing modification comprises a 2'-NH₂ nucleotide. In certain embodiments, the stability-enhancing modification comprises a 2'-H (or 2'-deoxy) nucleotide. In certain embodiments, the stability-enhancing modification comprises a 2'-arabino or a 2'-F-arabino. In certain embodiments, the stability-enhancing modification comprises a 4'-thioribosyl sugar moiety. In certain embodiments, the stability-enhancing modification comprises a 3'-phosphorothioate group. In certain embodiments, the stability-enhancing modification comprises a 3'-phosphonoacetate group. In certain embodiments, the stability-enhancing modification comprises a nucleotide containing a 3'-thiophosphonoacetate group. In certain embodiments, the stability-enhancing modification comprises a nucleotide containing a 3'-methylphosphonate group. In certain embodiments, the stability-enhancing modification comprises a nucleotide containing a 3'-boranophosphate group. In certain embodiments, the stability-enhancing modification comprises a nucleotide containing a 3'-phosphorodithioate group. In certain embodiments, the stability-enhancing modification comprises a locked nucleic acid ("LNA") nucleotide. In certain, embodiments, the stability-enhancing modification comprises an unlocked nucleic acid ("ULNA") nucleotide.

In certain embodiments, the stability-enhancing modification comprises a 2'-O-methyl and a 3'-phosphorothioate group on the same nucleotide. In certain embodiments, the stability-enhancing modification comprises a 2'-O-methyl and a 3'-phosphonoacetate group on the same nucleotide. In certain embodiments, the stability-enhancing modification comprises a 2'-O-methyl and a 3'-thiophosphonoacetate group on the same nucleotide. In certain embodiments, the stability-enhancing modification comprises a 2'-fluoro and a 3'-phosphonoacetate group on the same nucleotide. In certain embodiments, the stability-enhancing modification comprises a 2'-fluoro and a 3'-thiophosphonoacetate group on the same nucleotide.

32

In certain embodiments, the modification is a specificity-altering modification. In some embodiments, specificity enhancement may be achieved by enhancing on-target binding and/or cleavage, or reducing off-target binding and/or cleavage, or a combination of both. In some other embodiments, specificity reduction may be achieved, for example, by reducing on-target binding and/or cleavage, or increasing off-target binding and/or cleavage, or a combination of both.

In certain embodiments, the specificity-altering modification comprises a 2'-O-methyl. In certain embodiments the specificity-altering modification comprises a 2'-halo, such as 2'-fluoro.

In certain embodiments, the specificity-altering modification comprises a 2-thiouracil base (2-thioU). In certain embodiments, the specificity-altering modification comprises 2-thioC. In certain embodiments, the specificity-altering modification comprises 4-thioU. In certain embodiments, the specificity-altering modification comprises 6-thioG. In certain embodiments, the specificity-altering modification comprises 2-aminoA. In certain embodiments, the specificity-altering modification comprises a 2-aminopurine. In certain embodiments, the specificity-altering modification comprises pseudouracil. In certain embodiments, the specificity-altering modification comprises hypoxanthine. In certain embodiments, the specificity-altering modification comprises 7-deazaguanine. In certain embodiments, the specificity-altering modification comprises 7-deaza-8-azaguanine. In certain embodiments, the specificity-altering modification comprises 7-deazaadenine. In certain embodiments, the specificity-altering modification comprises 7-deaza-8-azaadenine. In certain embodiments, the specificity-altering modification comprises 5-methylC. In certain embodiments, the specificity-altering modification comprises 5-methylIU. In certain embodiments, the specificity-altering modification comprises 5-hydroxymethylcytosine. In certain embodiments, the specificity-altering modification comprises 5-hydroxymethyluracil. In certain embodiments, the specificity-altering modification, comprises 5,6-dehydrouracil. In certain embodiments, the specificity altering modification comprises 5-propynylcytosine. In certain embodiments, the specificity-altering modification comprises 5-propynyluracil. In certain embodiments, the specificity-altering modification comprises 5-ethynylcytosine. In certain embodiments, the specificity-altering modification comprises 5-ethynyluracil. In certain embodiments, the specificity-altering modification comprises 5-allylIU. In certain embodiments, the specificity-altering modification comprises 5-allylC. In certain embodiments, the specificity-altering modification comprises 5-aminoallylU. In certain embodiments, the specificity-altering modification comprises 5-aminoallylC. In certain embodiments, the specificity-altering modification comprises an abasic nucleotide. In certain embodiments, the specificity-altering modification comprises a Z base. In certain embodiments, the specificity-altering modification comprises P base. In certain embodiments, the specificity-altering modification comprises a UNA base. In certain embodiments, the specificity-altering modification comprises isoC. In certain embodiments, the specificity-altering modification comprises isoG. In certain embodiments, the specificity-altering modification comprises 5-methyl-pyrimidine. In certain embodiments, the specificity-altering modification comprises x(A,G,C,T). In certain embodiments, the specificity-altering modification comprises y(A,G,C,T).

In certain embodiments, the specificity-altering modification comprises a phosphorothioate internucleotide linkage. In certain embodiments, the specificity-altering modifi-

US 10,337,001 B2

33

fication comprises a phosphonoacetate internucleotide linkage. In certain embodiments, the specificity-altering modification comprises a thiophosphonoacetate internucleotide linkage. In certain embodiments, the specificity-altering modification comprises a methylphosphonate internucleotide linkage. In certain embodiments, the specificity-altering modification comprises a boranophosphate internucleotide linkage. In certain embodiments, the specificity-altering modification comprises a phosphorodithioate internucleotide linkage. In certain embodiments, the specificity-altering modification comprises a ULNA. In certain embodiments, the specificity-altering modification comprises an LNA.

In certain embodiments, the modification alters RNA base pairing by, for example, altering the melting temperature (T_m) of the guide RNA relative to a guide RNA without the modification. In certain embodiments, the modification lowers the T_m of the guide RNA relative to a guide RNA without the modification. In certain embodiments, the modification raises the T_m of the guide RNA relative to a guide RNA without the modification.

In certain embodiments, the specificity-altering modification lowers the T_m of a base pairing interaction. In certain embodiments, the modification that lowers the T_m of the base pairing interaction is a 2'-deoxy, as it is well-known in the art that DNA/DNA base pairs have lower T_m than their respective counterpart in RNA/DNA duplexes. In certain embodiments, the modification that lowers the T_m of the base pairing interaction is 2-thiouracil, which slightly lowers T_m of G-U wobble pair. In certain embodiments, the modification that lowers the T_m of the base pairing interaction is a phosphorothioate internucleotide linkage or a phosphorodithioate internucleotide linkage, which lower the T_m by $\sim 0.5^\circ$ C. per modification. In certain embodiments, the modification that lowers the T_m of the base pairing interaction is a phosphonoacetate internucleotide linkage, which lowers the T_m by $\sim 0.5\text{--}0.8^\circ$ C. per modification. In certain embodiments, the modification that lowers the T_m of the base pairing interaction is a phosphorothioate internucleotide linkage, which lowers the T_m by $\sim 1.3^\circ$ C. per modification. In certain embodiments, the modification that lowers the T_m of the base pairing interaction is unlocked nucleic acid ("ULNA"), which lowers the T_m by $\sim 5\text{--}8^\circ$ C. per modification. In certain embodiments, the modification that lowers the T_m of the base pairing interaction is 2'-O-methyl-3'-methylphosphonate.

In certain embodiments, the specificity-altering modification raises the T_m of a base pairing interaction. In certain embodiments, the modification that raises the T_m of the base pairing interaction is a 2'-O-methyl, which raises T_m by $\sim 0.5\text{--}0.7^\circ$ C. per modification. In certain embodiments, the modification that raises the T_m of the base pairing interaction is a 2'-F, which raises T_m by $\sim 1^\circ$ C. per modification. In certain embodiments, the modification that raises the T_m of the base pairing interaction is a 2-thiouracil, which raises T_m of A-U pair (and, as noted above, slightly lowers T_m of G-U wobble pair). In certain embodiments, the modification that raises the T_m of the base pairing interaction is a 4-thiouracil, which raises T_m of G-U wobble pair and slightly raises T_m of A-U pair. In certain embodiments, the modification that raises the T_m of the base pairing interaction is a 2-amino-adenine, which raises T_m of its base pairing with U by $\sim 1^\circ$ C. per modification. In certain embodiments, the modification that raises the T_m of the base pairing interaction is a 5-methyl-uracil (5-methylU) (see, e.g., Wang & Kool (1995) *Biochemistry*, 34, 4125-32). In certain embodiments, the modification that raises the T_m of

34

the base pairing interaction is a 5-methyl-cytosine (5-methylC). In certain embodiments, the modification that raises the T_m of the base pairing interaction is a locked nucleic acid ("LNA"), which raises T_m by 2-10° C. per modification.

In certain embodiments/the modification alters transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, the modification increases transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, the modification decreases transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, the modification neutralizes the anionic charge on phosphate to allow passive diffusion into cells. In certain embodiments, the charge-neutralizing modification comprises a phosphonoacetate alkyl ester internucleotide linkage, such as a phosphonoacetate methyl ester internucleotide linkage.

In certain embodiments, the modification alters the immunostimulatory effect of the guide RNA relative to a guide RNA without the modification. It was initially discovered that unmethylated bacterial DNA and synthetic analogs thereof are ligands for TLR9 (see Hemmi et al. (2000) *Nature*, 408, 740-5). The stimulation of TLR9 can be mitigated in the dinucleotide motif for example by modifying the C and G residues. The use of 5-methylcytosine, 2-aminocytosine, 2-thiocytosine, 5-methylisocytosine, P nucleobase (6-(β -D-2'-Deoxyribofuranosyl)-3,4-dihydro-8H-pyrimido[4,5-c][1,2]oxazin-7-one), and 2'-O-methylcytosine all result in loss or decrease in TLR9 stimulation. In certain embodiments, use of 6-thioguanine, 2,6-diaminopurine, 2-aminopurine, xanthosine, inosine, 7-deazaxanthosine, isoguanine, 8-oxoguanine, nebularine, 8-bromoguanine, K-nucleobase (2-amino-N-methoxyadenosine), and/or 2'-O-methylguanine can result in loss or decrease in TLR9 stimulation. In some embodiments, use of phosphodiester modifications can lower or eliminate the TLR9 response. Typically, synthetically incorporated phosphorothioates can decrease the TLR9 response to a limited extent, as is thought to result from the presence of two stereoisomers of each phosphorothioate in synthetic RNA. However, it has been shown that phosphorothioate-modified DNA lacking CpG motifs stimulate TLR9 to a rather small extent. The negative charge on the phosphorus is an important element for recognition by TLR9 and therefore removing the negative charge using alkylphosphonate can result in loss or decrease in TLR9 stimulation. The use of phosphonoacetate (PACE) internucleotide linkages between deoxynucleotides in 5' and 3' terminal sequences can significantly increase the TLR9 response; however, the use of thiophosphonoacetate (thio-PACE) internucleotide linkages between deoxynucleotides in 5' and 3' terminal sequences can result in loss or decrease in TLR9 stimulation. In certain embodiments, use of sugar modifications that favor C3'-endo conformation such as 2'-O-methyl modifications can be incorporated at 5' and 3' termini to decrease the TLR9 response. TLR 7 and TLR8 can be stimulated by molecules containing 7-deazaguanine and by single-stranded RNA (see e.g., Heil et al. (2004) *Science*, 303, 1526-9). TLR3 has been implicated in cellular immunoresponses to virus-derived double-stranded RNA. In certain embodiments, these TLR responses can be mitigated for example by using 2'-O-methyl modifications, modified phosphodiester linkages containing sulfur, or modifications that decrease internucleotide negative charge such as methylphosphonate and/or phosphonoacetate internucleotide linkages.

US 10,337,001 B2

35

In certain embodiments, the modification enhances stability and specificity of the guide RNA relative to a guide RNA without the modification. In certain embodiments, the modification enhances stability and transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, the modification enhances specificity and transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, the modification enhances the overall efficacy of the guide RNA relative to a guide RNA without the modification.

C. Guide RNA with a Combination of Modifications

In one aspect, the present technology provides a guide RNA having a combination of two or more modifications.

In certain embodiments, the two modifications are on the same nucleotide (for example, one nucleotide comprises a 2'-O-methyl and a 3'-thiophosphonoacetate moiety). In other embodiments, the two modifications are on two different nucleotides (for example, one nucleotide has a 2-thioU base and another nucleotide has a 2'-O-methyl group).

In certain embodiments, each modification in the guide RNA is the same. In certain embodiments, at least one modification in the guide RNA is different from at least one other modification in the guide RNA. In certain embodiments, a single nucleotide within the guide RNA possesses two or more modifications.

In certain embodiments, the guide RNA comprises a combination of different types of modifications, and at least one type in the combination exists in multiple places in the guide RNA. In certain embodiments, at least one type in the combination appears 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 times in the guide RNA.

In certain embodiments, at least one type of the modifications in the combination appears in two or more modified nucleotides. In certain embodiments, at least one type of the modifications in the combination appears in three or more modified nucleotides. In certain embodiments, the modified nucleotides are not arranged contiguously in the sequence, or at least not entirely, as one or more unmodified nucleotides may intercede. In certain embodiments, the modified nucleotides are arranged contiguously. In certain embodiments, the guide RNA comprises a stretch of contiguous modified nucleotides of the same type. In certain embodiments, the stretch has at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 modified nucleotides.

In certain embodiments, at least one type of the modifications in the combination is within the 5' portion of the guide RNA. In certain embodiments, at least one type of the modifications in the combination is within the first five (5) nucleotides of the 5' portion of the guide RNA. In certain embodiments, at least one type of the modifications in the combination is within the first three (3) nucleotides of the 5' portion of the guide RNA. In certain embodiments, at least one type of the modifications in the combination is within the 3' portion of the guide RNA. In certain embodiments, at least one type of the modifications in the combination is within the last five (5) nucleotides of the 3' portion of the guide RNA. In certain embodiments, at least one type of the modifications in the combination is within the last three (3) nucleotides of the 3' portion of the guide RNA. In certain embodiments, at least one type of the modifications in the

36

conciliation is within the internal region (i.e., between the 5' end and the 3' end) of the guide RNA.

In certain embodiments, at least one type of the modifications in the combination is incorporated in the 5' portion or 5 3' portion of the guide RNA, particularly within the first 5 or 10 nucleotides of the 5' portion or within the last 5 or 10 nucleotides of the 3' portion to, for example, protect the RNA from degradation by nucleases or for other purposes. In certain embodiments, at least one type of the modifications in the combination is in the 5' portion and at least one type of the modifications in the combination is in the 3' portion of the guide RNA, particularly within the first 5 or 10 nucleotides of the 5' portion and within the last 5 or 10 nucleotides of the 3' portion to, for example, protect the RNA from degradation by nucleases or for other purposes. In certain embodiments, a guide RNA comprises 20 or fewer, alternatively 15 or fewer, alternatively 15 or fewer, alternatively 10 or fewer, alternatively 5 or fewer, alternatively 3 or fewer deoxyribonucleotide residues in the 5' portion of the guide RNA.

In certain embodiments, at least one type of the modifications in the combination is within the crRNA segment of the guide RNA. In certain embodiments, at least one type of the modifications in the combination is within the guide sequence of the crRNA. In certain embodiments, at least one type of the modifications in the combination is within the first five (5) nucleotides of the crRNA segment. In certain embodiments, at least one type of the modifications in the combination is within the first three (3) nucleotides of the crRNA segment. In certain embodiments, at least one type of the modifications in the combination is within the tracrRNA segment of the guide RNA. In certain embodiments, at least one type of the modifications in the combination is within the last five (5) nucleotides of the tracrRNA segment of the guide RNA. In certain embodiments, at least one type of the modifications in the combination is within the last three (3) nucleotides of the tracrRNA segment of the guide RNA.

40 In certain embodiments, a first type of modification in the combination's within the 5' portion of the guide RNA and a second type of modification in the combination is within the internal region (i.e., between the 5' end and the 3' end) of the guide RNA. In certain embodiments, the first type of modification is within the first five (5) nucleotides of the 5' portion of the guide RNA. In certain embodiments, the first type of modification is within the first three (3) nucleotides of the 5' portion of the guide RNA.

45 In certain embodiments, a first type of modification in the combination is within the internal region (i.e., between the 5' end and the 3' end) of the guide RNA and a second type of modification in the combination is within the 3' portion of the guide RNA. In certain embodiments, the second type of modification is within the last five (5) nucleotides of the 3' portion of the guide RNA. In certain embodiments, the second type of modification is within the last three (3) nucleotides of the 3' portion of the guide RNA.

50 In certain embodiments, a first type of modification in the combination is within the 5' portion of the guide RNA and a second type of modification in the combination is within the 3' portion of the guide RNA. In certain embodiments, the first type of modification is within the first five (5) nucleotides of the 5' portion of the guide RNA. In certain embodiments, the first type of modification is within the first three (3) nucleotides of the 5' portion of the guide RNA. In certain embodiments, the second type of modification is within the last five (5) nucleotides of the 3' portion of the guide RNA.

US 10,337,001 B2

37

In certain embodiments, the second type of modification is within the last three (3) nucleotides of the 3' portion of the guide RNA.

In certain embodiments, a first type of modification in the combination is within the 5' portion of the guide RNA, a second type of modification in the combination is within the internal region (i.e., between the 5' end and the 3' end) of the guide RNA, and a third type of modification in the combination is within the 3' portion of the guide RNA. In certain embodiments, the first type of modification is within the first five (5) nucleotides of the 5' portion of the guide RNA. In certain embodiments, the first type of modification is within the first three (3) nucleotides of the 5' portion of the guide RNA. In certain embodiments, the third type of modification is within the last five (5) nucleotides of the 3' portion of the guide RNA. In certain embodiments, the third type of modification is within the last three (3) nucleotides of the 3' portion of the guide RNA.

In certain embodiments, a first type of modification in the combination is within the crRNA segment of the guide RNA and a second type of modification in the combination is within the trace segment. In certain embodiments, the first type of modification is within the guide sequence of the crRNA. In certain embodiments, the first type of modification is within the first five (5) nucleotides of the crRNA segment. In certain embodiments, the first type of modification is within the first three (3) nucleotides of the crRNA segment. In certain embodiments, the second type of modification is within the last five (5) nucleotides of the tracrRNA segment of the guide RNA. In certain embodiments, the second type of modification is within the last three (3) nucleotides of the tracrRNA segment of the guide RNA.

In certain embodiments, a first type and a second type of modification in the combination are within the crRNA segment of the guide RNA. In certain embodiments, the first type of modification is within the guide sequence of the crRNA. In certain embodiments, the first type of modification is within the first five (5) nucleotides of the crRNA segment. In certain embodiments, the first type of modification is within the first three (3) nucleotides of the crRNA segment.

In certain embodiments, a first type and a second type of modification in the combination are within the crRNA segment of the guide RNA and a third type of modification in the combination is within the trace segment. In certain embodiments, the first type of modification is within the guide sequence of the crRNA. In certain embodiments, the first type of modification is within the first five (5) nucleotides of the crRNA segment. In certain embodiments, the first type of modification is within the first three (3) nucleotides of the crRNA segment. In certain embodiments, the third type of modification is within the last five (5) nucleotides of the tracrRNA segment of the guide RNA. In certain embodiments, the third type of modification is within the last three (3) nucleotides of the tracrRNA segment of the guide RNA.

In certain embodiments, at least one of the modifications in the combination comprises an end modification, such as a 5' end modification or a 3' end modification as described above. In certain embodiments, the end modification comprises dimethoxytrityl.

In certain embodiments, at least one of the modifications in the combination comprises a modified base. In certain embodiments, the modified gRNA comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 modified bases. In other embodiments, the modified

38

gRNA comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified bases. In certain embodiments, all bases in a gRNA are modified.

In certain embodiments, at least one of the modifications in the combination comprises a modified sugar. In certain embodiments, the modified gRNA comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 modified sugars. In other embodiments, the modified gRNA comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified sugars. In certain embodiments, all sugars in a gRNA are modified.

In certain embodiments, at least one of the modifications in the combination comprises a modified backbone (i.e., an internucleotide linkage other than a natural phosphodiester). In certain embodiments, the modified-gRNA comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 modified internucleotide linkages. In other embodiments, the modified gRNA comprises at least 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130 or 140 modified internucleotide linkages. In certain embodiments, all internucleotide linkages in a gRNA are modified.

In certain embodiments, at least one of the modifications in the combination comprises a 2'-O-methyl, a 2'-fluoro, a 2'-amino, a 2'-deoxy, a 2'-arabino, a 2'-F-arabino, a 2-thiouracil, a 2-aminoadenine, a 5-methylcytosine, a 5-aminoallyluracil, a Z base, a 3'-phosphorothioate, a 3'-phosphonacetate, a 3'-phosphonoacetate ester, a 3'-thiophosphonoacetate, a 3'-thiophosphonoacetate ester, a 3'-methylphosphonate, a 3'-boranophosphonate, a 3'-phosphorodithioate, or combinations thereof. In certain embodiments, at least one of the modifications in the combination comprises a 2'-O-methyl, a 2'-deoxy, a Z base, a phosphorothioate internucleotide linkage, a phosphonoacetate internucleotide linkage, a thiophosphonoacetate internucleotide linkage, or combinations thereof. In certain embodiments, at least one of the modifications in the combination comprises a 2'-F, a 2-thioU, a 4-thioU, a 2-aminoA, a 5-methylC, a 5-methylU, a 5-aminoallylU, or combinations thereof. In certain embodiments, at least one of the modifications in the combination is an "end" modification such as terminal phosphate, a PEG, a terminal amine, a terminal linker such as a hydrocarbon linker, a substituted hydrocarbon linker, a square linker, a triazolo linker, an internal linker such as 2-(4-butylamidofluorescein)propane-1,3-diol bis(phosphodiester) linker, a linker conjugated to a dye, a linker conjugated to a non-fluorescent label, a linker conjugated to a tag or a linker conjugated to a solid support such as for example a bead or microarray. In certain embodiments, at least two of the modifications in the combination comprise a 2'-O-methyl nucleotide and phosphorothioate internucleotide linkage, a 2'-O-methyl nucleotide and phosphonoacetate internucleotide linkage, or a 2'-O-methyl nucleotide and thiophosphonoacetate internucleotide linkage. In certain embodiments, at least two of the modifications in the combination comprise a 2'-O-methyl nucleotide and phosphonocarboxylate internucleotide linkage, a 2'-O-methyl nucleotide and phosphonocarboxylate ester internucleotide linkage, a 2'-O-methyl nucleotide and thiophosphonocarboxylate linkage, a 2'-O-methyl nucleotide and thiophosphonocarboxylate ester internucleotide linkage, or combinations thereof. In other embodiments, the modifications in the combination further comprise a 2-thiouracil, 2-thiocytosine, 4-thiouracil, 6-thioguanine, 2-aminoadenine, 2-aminopurine, pseudouracil, inosine, 7-deazaguanine, 7-deaza-8-azaguanine, 7-deazaadenine, 7-deaza-8-azaade-

25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
110
120
130
140

US 10,337,001 B2

39

nine, 5-methylcytosine, 5-methyluracil, 5-hydroxymethylcytosine, 5-hydroxymethyluracil, 5,6-dehydrouracil, 5-propynylcytosine, 5-propynyluracil, 5-ethynylcytosine, 5-ethynyluracil, 5-allyluracil, 5-allylcytosine, 5-aminoallyluracil, 5-aminoallyl-cytosine, or an abasic nucleotide.

In certain embodiments, at least one of the modifications in the combination comprises a 2'-O-methyl-3'-phosphorothioate. In certain embodiments, at least one of the modifications in the combination comprises a 2'-O-methyl-3'-phosphonoacetate. In certain embodiments, at least one of the modifications in the combination comprises a 2'-O-methyl-3'-thiophosphonoacetate. In certain embodiments, at least one of the modifications in the combination comprises a 2'-halo-3'-phosphorothioate. In certain embodiments, at least one of the modifications in the combination comprises a 2'-halo-3'-phosphonoacetate. In certain embodiments, at least one of the modifications in the combination comprises a 2'-halo-3'-thiophosphonoacetate. In certain embodiments, at least one of the modifications in the combination comprises a 2'-fluoro-3'-phosphorothioate. In certain embodiments, at least one of the modifications in the combination comprises a 2'-fluoro-3'-phosphonoacetate. In certain embodiments, at least one of the modifications in the combination comprises a 2'-fluoro-3'-thiophosphonoacetate. Possible combinations of at least two or three modifications are represented in FIG. 6 and FIG. 7 respectively and are incorporated herein by reference.

In certain embodiments, the guide RNA comprises an oligonucleotide represented by Formula (III) or Formula (IV):

W-Y-Q	(III); or
Y-W-X-Q	(IV)

wherein Q and W each independently represent a nucleotide or a stretch of nucleotides of the oligonucleotide comprising at least one modification and Y and X each independently represent an unmodified portion of the oligonucleotide.

In certain embodiments, W is within the 5' portion of the guide RNA. In certain embodiments, W is at least partially within the first five (5) nucleotides of the 5' portion of the guide RNA. In certain embodiments, W is at least partially within the first three (3) nucleotides of the 5' portion of the guide RNA. In certain embodiments, W is within the internal region (i.e., between the 5' end and the 3' end) of the guide RNA.

In certain embodiments, Q is within the 3' portion of the guide RNA. In certain embodiments, Q is at least partially within the last five (5) nucleotides of the 3' portion of the guide RNA. In certain embodiments, Q is at least partially within the last three (3) nucleotides of the 3' portion of the guide RNA. In certain embodiments, Q is within the internal region (i.e., between the 5' end and the 3' end) of the guide RNA.

In certain embodiments, W comprises an end modification as described above, such as a 5' end or a 3' end modification. In certain embodiments, the end modification comprises dimethoxytrityl.

In certain embodiments, at least one of W or Q comprises a modified base as described above. In certain embodiments, at least one of W or Q comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 modified bases. In certain embodiments, at least one of W or Q comprises more than one, such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 modified sugars.

40

one modified base, such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 modified bases.

In certain embodiments, at least one of W or Q comprises a modified sugar as described above. In certain embodiments, at least one of W or Q comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 modified sugars. In certain embodiments, at least one of W or Q comprises more than one, such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 modified sugars.

In certain embodiments, at least one of W or Q comprises a modified backbone (i.e., an internucleotide linkage other than a phosphodiester) as described above. In certain embodiments, at least one of W or Q comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 modified internucleotide linkages. In certain embodiments, at least one of W or Q comprises more than one, such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 modified internucleotide linkages.

In certain embodiments, at least one of W or Q comprises a 2'-O-methyl nucleotide, a 2'-F nucleotide, a 2'-amino nucleotide, a 2'-deoxy nucleotide, a 2-thiouridine nucleotide, a 2-aminoadenine nucleotide, a 6-thioguanosine nucleotide, a 5-methylcytidine nucleotide, a 5-aminoallyluridine nucleotide, a Z nucleotide, a 3'-phosphorothioate internucleotide linkage, a 3'-phosphonoacetate internucleotide linkage, a 3'-phosphonoacetate ester internucleotide linkage, a 3'-thiophosphonoacetate internucleotide linkage, a 3-thiophosphonoacetate ester internucleotide linkage, a 3'-methylphosphonate internucleotide linkage, a 3'-boranophosphate internucleotide linkage, a 3'-phosphorodithioate internucleotide linkage, or combinations thereof.

In certain embodiments, at least one of W or Q comprises a 2'-O-methyl and a 3'-phosphorothioate group on the same nucleotide. In certain embodiments, at least one of W or Q comprises a 2'-O-methyl and a 3'-phosphonoacetate group linkage on the same nucleotide. In certain embodiments, at least one of W or Q comprises a 2'-O-methyl and 3'-thiophosphonoacetate group on the same nucleotide. In certain embodiments, at least one of W or Q comprises a 2'-F and a 3'-phosphorothioate group on the same nucleotide. In certain embodiments, at least one of W or Q comprises a 2'-F and a 3'-phosphonoacetate group linkage on the same nucleotide. In certain embodiments, at least one of W or Q comprises a 2'-F and 3'-thiophosphonoacetate group on the same nucleotide.

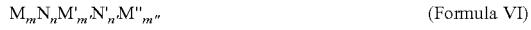
In certain embodiments, at least one of W or Q comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 modified nucleotides. In certain embodiments, each of the modified nucleotides within at least one of W or Q comprises the same modification or modifications. In certain embodiments, W comprises a modified nucleotide that is different than a modified nucleotide in Q. In certain embodiments, at least one of W or Q comprises two or more modified nucleotides. In certain embodiments, at least one of W or Q comprises three or more modified nucleotides. In certain embodiments, the modified nucleotides are not arranged contiguously in the sequence, or at least not entirely, as one or more unmodified nucleotides may intercede. In certain embodiments, the modified nucleotides are arranged contiguously. In certain embodiments, at least one of W or Q comprises at least one-contiguous stretch of modified nucleotides. In certain embodiments, at least one of

US 10,337,001 B2

41

W or Q comprises a contiguous stretch of at least three (3) modified nucleotides. In certain embodiments, at least one of W or Q comprises a contiguous stretch of at least four (4) modified nucleotides. In certain embodiments, at least one of W or Q comprises a contiguous stretch of at least five (5) modified nucleotides.

In certain embodiments, the guide RNA comprises a nucleotide sequence of Formula (V) or Formula (VI):



wherein each M independently represents a modified ribonucleotide;

wherein each N independently represents an unmodified ribonucleotide;

wherein each M' independently represents a modified ribonucleotide;

wherein each N' independently represents an unmodified ribonucleotide;

wherein each M'' independently represents a modified ribonucleotide;

wherein m is an integer between 0 and 40, n is an integer between 0 and 130, m' is an integer between 0 and 10, n' is an integer between 0 and 130, m'' is an integer between 0 and 10, provided that $m+m'+m''$ is greater than or equal to 1 and $50 < m+n+m''+n''+m'' \leq 150$.

In certain embodiments, each M is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, and a 2'-deoxynucleotide. In certain embodiments, each M is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, and a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M is independently selected from the group consisting of a 2'-O-methyl-3'-PACE ribonucleotide and a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, where $m>1$, any given M is the same or different from any other M. In certain embodiments, where $m>1$, each M comprises the same modification or modification.

In certain embodiments, each M' is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, and a 2'-deoxynucleotide. In certain embodiments, each M' is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, and a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M' is independently selected from the group consisting of a 2'-O-methyl-3'-PACE ribonucleotide and a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, where $m'>1$, any given M' is the same or different from any other M'. In certain embodiments, where $m'>1$, each M' comprises the same modification or modification.

In certain embodiments, each M'' is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, and a 2'-deoxynucleotide. In certain embodiments, each M'' is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, and a 2'-O-methyl-3'-thioPACE ribonucleotide.

42

otide, and a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M'' is independently selected from the group consisting of a 2'-O-methyl-3'-PACE ribonucleotide and a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, where $m''>1$, any given M'' is the same or different from any other M''. In certain embodiments, where $m''>1$, each M'' comprises the same modification or modifications.

In certain embodiments, each M is a 2'-O-methyl-3'-PACE ribonucleotide; m is selected from an integer between 1 and 10; each N is independently selected from the group consisting of A, U, C, and G; n is selected from an integer between 10 and 130; each M' is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S)ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, a 2'-deoxynucleotide, and a Z nucleotide; m' is selected from an integer between 1 and 10; each N is independently selected from the group consisting of A, U, C, and G; and n' is selected from an integer between 0 and 130. In certain embodiments, each M' is a 2'-O-methyl-3'-PACE ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl-3'-P(S) ribonucleotide. In certain embodiments, each M' is a Z nucleotide.

In certain embodiments, each M is a 2'-O-methyl-3'-thioPACE ribonucleotide; m is selected from an integer between 1 and 10; each N is independently selected from the group consisting of A, U, C, and G; n is selected from an integer between 10 and 130; each M' is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, a 2'-deoxynucleotide, and a Z nucleotide; m' is selected from an integer between 1 and 10; each N is independently selected from the group consisting of A, U, C, and G; and n' is selected from an integer between 0 and 130. In certain embodiments, each M' is a 2'-O-methyl-3'-PACE ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl-3'-P(S) ribonucleotide. In certain embodiments, each M' is a Z nucleotide.

In certain embodiments, each M is a 2'-O-methyl-3'-P(S) ribonucleotide; m is selected from an integer between 1 and 10; each N is independently selected from the group consisting of A, U, C, and G; n is selected from an integer between 10 and 130; each M' is independently selected from the group consisting of a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, a 2'-deoxynucleotide, and a Z nucleotide; m' is selected from an integer between 1 and 10; each N is independently selected from the group consisting of A, U, C, and G; and n' is selected from an integer between 0 and 130. In certain embodiments, each M' is a 2'-O-methyl-3'-PACE ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl ribonucleotide. In certain embodiments, each M' is a 2'-O-methyl-3'-P(S) ribonucleotide. In certain embodiments, each M' is a Z nucleotide.

In certain embodiments, each M is independently selected from the group consisting of a 2'-O-methylribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, a 2'-deoxynucleotide, and a Z nucleotide; m is selected from

US 10,337,001 B2

43

an integer between 0 and 10; each N is independently selected from the group consisting of A, U, C, and G; n is selected from an integer between 10 and 15; each M' is a 2'-O-methyl ribonucleotide; m' is selected from an integer between 1 and 5; each N is independently selected from the group consisting of A, U, C, and G; and n' is selected from an integer between 0 and 130. In certain embodiments, each M is a 2'-O-methyl-3'-PACE ribonucleotide. In certain embodiments, each M is a 2'-O-methyl-3'-thioPACE ribonucleotide. In certain embodiments, each M is a 2'-O-methyl ribonucleotide. In certain embodiments, each M is a 2'-O-methyl-3'-P(S)ribonucleotide. In certain embodiments, m is 0; n is selected from an integer between 10 and 15, m' is selected from an integer between 1 and 5; and n' is selected from an integer between 0 and 130.

In certain embodiments, at least one of the modifications in the combination is a stability-altering modification. In certain embodiments, at least one of the modifications in the combination increases nuclease resistance of the guide RNA relative to a guide RNA without the modification, thus it enhances the stability of the guide RNA.

In certain embodiments, at least one of the modifications in the combination is a stability-enhancing modification as described above.

In certain embodiments, at least one of the modifications in the combination is a specificity-altering modification as described above.

In certain embodiments, at least one of the modifications in the combination alters RNA base pairing. In certain embodiments, at least one of the modifications in the combination lowers the Tm of a base pairing interaction as described above. In certain embodiments, at least one of the modifications in the combination raises the Tm of a base pairing interaction as described above.

In certain embodiments, at least one of the modifications in the combination alters transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, at least one of the modifications in the combination increases transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, at least one of the modifications in the combination decreases transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, at least one of the transfection-increasing modifications in the combination comprises a phosphonoacetate alkyl ester internucleotide linkage, such as a phosphonoacetate methyl ester internucleotide linkage.

In certain embodiments, at least one of the modifications in the combination enhances stability and specificity of the guide RNA relative to a guide RNA without the modification. In certain embodiments, at least one of the modifications in the combination enhances stability and transfection efficiency of the guide RNA relative to a guide RNA without the modification. In certain embodiments, at least one of the modifications in the combination enhances specificity and transfection efficiency of the guide RNA relative to a guide RNA without the modification.

In certain embodiments, at least one of the modifications in the combination alters the secondary structure of the guide RNA. This modification alters the base-pairing of any of the RNA/RNA internal duplexes in the guide RNA. Some of these modifications increase the base pairing of the RNA/RNA structure or alternatively increase the Tm of the RNA/RNA duplex, whereas other modifications decrease the base pairing (or Tm) of the RNA/RNA duplex or duplexes. Such modifications include base modified nucleotides, particularly UNA nucleotides such as the 2-thiouridine

44

and 2-aminoadenosine pair, the Z/P nucleotide pair, the isoC/isoG pair, the 6-thioG/5-methylpyrimidine pair, and nucleotides with modifications on the sugar or the internucleotide linkages as discussed before.

5 In certain embodiments, the combination includes at least one modification or a set of modifications that increases nucleases resistance (i.e., stability) with at least one modification or a set of modifications that increases specificity (i.e., reduces off-target effects). In certain embodiments, the 10 combination includes at least one modification or a set of modifications that increases nucleases resistance (i.e., stability) with at least one modification or a set of modifications that raises the Tm of some bases pairing in the guide RNA. In certain embodiments, the combination includes at least one modification or a set of modifications that increases nucleases resistance (i.e., stability) with at least one modification or a set of modifications that lowers the Tm of some bases pairing of the guide RNA. In certain embodiments, the 15 combination includes at least one modification or a set of modifications that increases nucleases resistance (i.e., stability), at least one modification or a set of modifications that increases the Tm of some bases paring in the guide RNA, and at least one modification or a set of modifications that decreases the Tm of some base paring elsewhere in the guide RNA. In certain embodiments the combination includes at 20 least one modification or a set of modifications that increases nuclease resistance (i.e., stability) and at least one modification or a set of modifications that increases the binding of the guide RNA to Cas protein. In certain embodiments, the combination includes at least one modification or a set of modifications that increases nuclease resistance (i.e., stability) and at least one modification or a set of modifications that decreases the binding of the guide RNA to Cas protein. In certain embodiments, the guide RNA comprises a combination of the different types of modifications.

D. Guide RNA Structure

In certain embodiments, the guide RNA is able to form a 40 complex with a CRISPR-associated-protein. In certain embodiments, the CRISPR-associated protein is provided by or is derived from a CRISPR-Cas type II system, which has an RNA-guided polynucleotide binding and/or nuclease activity. In certain embodiments, the CRISPR-associated 45 protein is Cas9, a Cas9 mutant, or a Cas9 variant. In certain embodiments, the CRISPR-associated protein is the Cas9 nuclease from *Streptococcus pyogenes*. In certain embodiments, the CRISPR-associated protein is the Cas9 nuclease from *Streptococcus thermophilus*. In certain embodiments, the CRISPR-associated protein is the Cas9 nuclease from *Staphylococcus aureus*. In certain embodiments, the synthetic guide RNA or a synthetic guide RNA:CRISPR-associated protein complex maintains functionality of natural 50 guide RNA or a complex that does not have modified nucleotides. In certain embodiments, the functionality includes binding a target polynucleotide. In certain embodiments, the functionality includes nicking a target polynucleotide. In certain embodiments, the functionality includes cleaving a target polynucleotide. In certain embodiments, the target polynucleotide is within a nucleic acid in vitro. In certain embodiments, the target polynucleotide is within the genome of a cell in vivo or in vitro (such as in cultured cells or cells isolated from an organism). In certain embodiments, the target polynucleotide is a protospacer in DNA.

55 In certain embodiments, the crRNA segment comprises from 25 to 80 nucleotides. In certain embodiments, the crRNA segment comprises a guide sequence that is capable

US 10,337,001 B2

45

of hybridizing to a target sequence. In certain embodiments, the guide sequence is complementary to the target sequence or a portion thereof. In certain embodiments, the guide sequence comprises from 15 to 30 nucleotides. In certain embodiments, the crRNA segment comprises a stem sequence. In certain embodiments, the stem sequence comprises from 10 to 50 nucleotides. In certain embodiments, the crRNA segment comprises a 5'-overhang sequence. In certain embodiments, the 5'-overhang sequence comprises from 1 to 10 nucleotides, alternatively 1 to 5 nucleotides, alternatively 1, 2 or 3 nucleotides. In certain embodiments, the crRNA comprises both (i) a guide sequence that is capable of hybridizing to a target sequence and (ii) a stem sequence. In certain embodiments, the crRNA comprises (i) a 5'-overhang sequence, (ii) a guide sequence that is capable of hybridizing to a target sequence, and (iii) a stem sequence. In certain embodiments wherein the crRNA segment comprises a stem sequence, the tracrRNA segment comprises a nucleotide sequence that is partially or completely complementary to the stem sequence of the crRNA segment. In certain embodiments, the tracrRNA segment comprises at least one more duplex structure.

In certain embodiments, the guide RNA is a single guide RNA. In certain embodiments, the guide RNA is a single guide RNA, wherein the crRNA segment and the tracrRNA segment are linked through a loop L. In certain embodiments, the loop L comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides. In certain embodiments, the loop L comprises a nucleotide sequence of GNRA, wherein N represents A, C, G, or U and R represents A or G. In certain embodiments, the loop L comprises a nucleotide sequence of GAAA. In certain embodiments, the guide RNA comprises more than one loop.

The guide RNA comprises a 5' portion (i.e., the 5' half) and a 3' portion (i.e., the 3' half). In certain embodiments, the crRNA segment is 5' (i.e., upstream) of the tracrRNA segment. In certain embodiments, the tracrRNA segment is 5' relative to the crRNA segment.

In certain embodiments, the guide RNA comprises at least two separate RNA strands, for example, a crRNA strand and a separate tracrRNA strand. See, for example, FIG. 5A. In certain embodiments, each of the strands is a synthetic strand comprising one or more modifications. In certain embodiments, at least one of the strands is a synthetic strand comprising one or more modifications. In certain embodiments, the strands function together to guide binding, nicking, or cleaving of a target polynucleotide by a Cas protein, such as Cas9. In certain embodiments, the crRNA sequence and the tracrRNA sequence are on separate stands and hybridize to each other via two complementary sequences to form a stem or duplex.

In certain embodiments, the guide RNA is a single guide RNA comprising a crRNA sequence and a tracrRNA sequence. See, for example, FIG. 5B. In certain embodiments, the crRNA sequence and the tracrRNA sequence are connected by a loop sequence or "loop." In certain embodiments, a single guide RNA comprises a 5' portion and a 3' portion, wherein the crRNA sequence is upstream of the tracrRNA sequence.

In certain embodiments, the total length of the two RNA pieces can be about 50-220 (e.g., about 55-200, 60-190, 60-180, 60-170, 60-160, 60-150, 60-140, 60-130, and 60-120) nucleotides in length, such as about 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, or 220 nucleotides in length. Similarly, the single guide RNA (e.g., FIG. 5B) can be about 50-220 (e.g., about 55-200, 60-190, 60-180, 60-170, 60-160, 60-150, 60-140, 60-130, and

46

60-120) nucleotides in length, such as about 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, or 220 nucleotides in length.

As shown in FIGS. 5A and 5B, the synthetic guide RNA comprises (i) a crRNA sequence that comprises (a) a guide sequence (e.g., segment G₁-G_n, where each G represents a nucleotide in the guide sequence) capable of hybridizing to a target sequence in a nucleic acid, (b) a first stem sequence (e.g., segment X₁-X_n, where each X represents a nucleotide in the first-stem sequence) capable of hybridizing partially or completely to a second stem sequence, and, optionally (c) a 5' overhang sequence (e.g., segment O₁-O_n, where each O represents a nucleotide in the overhang sequence), and (ii) a tracrRNA sequence that comprises the second stem sequence (e.g., segment Y₁-Y_n, where each Y represents a nucleotide in the second stem sequence). The tracrRNA sequence further comprises segment T₁-T_n, where each T represents a nucleotide in the tracrRNA sequence. The synthetic guide RNA shown in FIG. 5A includes one or more modifications. Likewise, the synthetic guide RNA shown in FIG. 5B includes one or more modifications. In certain embodiments, the modification is located at any point along the length of the crRNA, the tracrRNA, or the single guide RNA comprising a crRNA segment, a tracrRNA segment, and, optionally, a loop. In certain embodiments, any nucleotide represented by O, G, X, Y, or T in the synthetic guide RNA shown in FIGS. 5A and 5B may be a modified nucleotide. The guide RNA shown in FIG. 5B represents a single guide RNA (sgRNA) where the crRNA segment and the tracrRNA segment are connected by a loop having the sequence GNRA, wherein N represents A, C, G, or U, and R represents A or G.

In certain embodiments, the crRNA segment of the guide RNA is 25-70 (e.g., 30-60, 35-50, or 40-45) nucleotides in length. In certain embodiments, the guide sequence is 12-30 (e.g., 16-25, 17-20, or 15-18) nucleotides in length. In some embodiments, a 5' portion of the crRNA does not hybridize or only partially hybridizes with the target sequence. For example, there can be a 5'-overhang on the crRNA segment.

In certain embodiments, the single guide RNA comprises a central portion including the stem sequence of the crRNA segment, the stem sequence of the tracrRNA segment, and, optionally, a loop that covalently connects the crRNA segment to the tracrRNA segment. In certain embodiments, the central segment of the single guide RNA is 8-60 (e.g., 10-55, 10-50, or 20-40) nucleotides in length.

In certain embodiments, the tracrRNA segment of the guide RNA is 10-130 (e.g., 10-125, 10-100, 10-75, 10-50, or 10-25) nucleotides in length. In certain embodiments, the tracrRNA segment includes one or more hairpin or duplex structures in addition to any hairpin or duplex structure in the central segment.

In certain embodiments, the tracrRNA is truncated compared to a reference tracrRNA, such as a naturally existing mature tracrRNA. A range of lengths has been shown to function in both the separate type (FIG. 5A) and the chimeric sgRNA type (FIG. 5B). For example, in certain embodiments, tracrRNA may be truncated from its 3' end by at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75 or 80 nts. In certain embodiments, the tracrRNA molecule may be truncated from its 5' end by at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75 or 80 nts. In certain embodiments, the tracrRNA molecule may be truncated from both the 5' and 3' end, e.g., by at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 or 20 nts from the 5' end and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 or 40 nts from the 3' end. See, e.g., Jinek et al. (2012) *Science*, 337, 816-21;

US 10,337,001 B2

47

Mali et al. (2013) *Science*, 339:6121, 823-6; Cong et al. (2013) *Science*, 339:6121, 819-23; and Hwang et al. (2013) *Nat. Biotechnol.* 31:3, 227-9; Jinek et al. (2013) *eLife*, 2, e00471. In certain embodiments, the tracrRNA is untruncated.

In certain embodiments, the disclosed modifications we in the crRNA segment or the tracrRNA segment or both. In certain embodiments, the disclosed modifications are in the guide sequence of the crRNA segment. In certain embodiments, the disclosed modifications are in the stem sequence of the crRNA segment. In certain embodiments, the disclosed modifications are in the 5'-overhang sequence of the crRNA segment. In certain embodiments, the disclosed modifications are in the stem sequence of the tracrRNA segment. In certain embodiments, the disclosed modifications are in the loop sequence of the guide RNA. In certain embodiments, the disclosed modifications are in the 5' portion of the guide RNA. In certain embodiments, the disclosed modifications are in the 3' portion of the guide RNA. In certain embodiments, the disclosed modifications are in the 5' portion of the guide RNA and the 3' portion of the guide RNA.

E. Synthesis of Guide RNA

In certain embodiments, guide RNAs, including single guide RNAs (sgRNAs; see FIGS. 1 and 5B) are produced by chemical synthesis using the art of synthetic organic chemistry. A guide RNA that comprises any nucleotide other than the four predominant ribonucleotides, namely A, C, G, and U, whether unnatural or natural, such as a pseudouridine, inosine or a deoxynucleotide, possesses a chemical modification or substitution at the nucleotide which is chemically/structurally distinct from any of the four predominant nucleotides in RNAs.

The synthetic guide RNAs described herein can be chemically synthesized. For example, the synthetic guide RNAs can be synthesized using TC chemistry by the method described in Dellinger et al. (2011) *J. Am. Chem. Soc.*, 133, 11540, U.S. Pat. No. 8,202,983, and US Patent Application 2010/0076183A1, the contents of which are incorporated by reference in their entireties. “TC chemistry” refers to the composition and methods of using RNA monomeric nucleotide precursors protected on the 2'-hydroxyl moiety by a thionocarbamate protecting group, to synthesize unmodified RNA or modified RNA comprising one or more modified nucleotides. The ability to chemically synthesize relatively long RNAs (as long as 200 mers or more) using TC-RNA chemistry allows one to produce guide RNAs with special features capable of outperforming those enabled by the four predominant ribonucleotides (A, C, G and U). Some synthetic guide RNAs described herein can also be made using methods known in the art that include in vitro transcription and cell-based expression. For example, 2'-fluoro NTPs can be incorporated into synthetic guide RNAs produced by cell-based expression.

Synthesis of guide RNAs can also be accomplished by chemical or enzymatic synthesis of RNA sequences that are subsequently ligated together by enzymes, or chemically ligated by chemical ligation, including but not limited to cyanogen bromide chemistry, “click” chemistry as published by R. Kumar et al. (2007) *J. Am. Chem. Soc.*, 129, 6859-64, or square conjugation chemistry as described by K. Hill in WO2013176844 titled “Compositions and methods for conjugating oligonucleotides.”

As further described below, a guide RNA disclosed herein, including those comprising modified nucleotides

48

and/or modified internucleotide linkages, can be used to perform various CRISPR-mediated functions (including but not limited to editing genes, regulating gene expression, cleaving target sequences, and binding to target sequences) in vitro or in vivo, such as in cell-free assays, in intact cells, or in whole organisms. For in vitro or in vivo applications, the RNA can be delivered into cells or whole organisms in any manner known in the art.

Libraries and Arrays

10 In one aspect, the present invention provides a set or library of multiple guide RNAs. In certain embodiments, the library contains two or more guide RNAs disclosed herein. The library can contain from about 10 to about 10^7 individual members, e.g., about 10 to about 10^2 , about 10^2 to 15 about 10^3 , about 10^3 to about 10^5 , from about 10^5 to about 10^7 members. An individual member of the library differs from other members of the library at least in the guide sequence, i.e., the DNA targeting segment of the gRNA. On the other hand, in certain embodiments, each individual 20 member of a library can contain the same or substantially the same nucleotide sequence for the tracrRNA segment as all the other members of the library. In this way, the library can comprise members that target different polynucleotides or different sequences in one or more polynucleotides.

25 In certain embodiments, the library comprises at least 10^2 unique guide sequences. In certain embodiments, the library comprises at least 10^3 unique guide sequences. In certain embodiments, the library comprises at least 10^4 unique guide sequences. In certain embodiments, the library comprises at least 10^5 unique guide sequences. In certain embodiments, the library comprises at least 10^6 unique guide sequences. In certain embodiments, the library comprises at least 10^7 unique guide sequences. In certain embodiments, the library targets at least 10 different polynucleotides. In certain embodiments, the library targets at least 10^2 different polynucleotides. In certain embodiments, the library targets at least 10^3 different polynucleotides. In certain embodiments, the library targets at least 10^4 different polynucleotides. In certain embodiments, the library targets at least 10^5 different polynucleotides. In certain embodiments, the library targets at least 10^6 different polynucleotides. In certain embodiments, the library targets at least 10^7 different polynucleotides.

30 In certain embodiments, the library comprises a collection 45 of guide RNAs having the same sequence and the same, modifications in a progressively shifted window that moves across the sequence of the members in the library. In certain embodiments, the windows collectively cover the entire length of the RNA.

50 In certain embodiments, the library allows one to conduct high-throughput, multi-target genomic manipulations and analyses. In certain embodiments, only the DNA-targeting segments of the guide RNAs are varied, while the Cas protein-binding segment is the same. In certain embodiments, a first portion of the library comprises guide RNAs 55 possessing a Cas-binding segment that recognizes, binds and directs a particular Cas protein and a second portion of the library comprises a different Cas-binding segment that recognizes, binds and directs a different Cas protein (e.g., a Cas protein from a different species), thereby allowing the library to function with two or more orthogonal Cas proteins. In certain embodiments, induced expression of a first orthogonal Cas protein utilizes the portion of the library which interacts with the first orthogonal Cas protein. In 60 certain embodiments, induced expression of a first and second orthogonal Cas protein utilizes the portions of the library which interact with the first and second orthogonal 65

US 10,337,001 B2

49

Cas proteins, respectively. In certain embodiments, induced expression of the first, and second orthogonal Cas proteins occur at different times. Accordingly, one can carry out large-scale gene editing or gene regulation by specifically manipulating or modifying multiple targets as specified in the library.

In certain embodiments, the library is an “arrayed” library, namely a collection of different features or pools of features in an addressable arrangement. For example, features of an array can be selectively cleaved and transferred to a microtiter plate such that each well in the plate contains a known feature or a known pool of features. In some other embodiments, the library is synthesized in a 48-columns or in a 96-columns microtiter plate format or in a 384-columns plate.

In certain embodiments, synthesis of the guide RNA of this invention may be conducted on a solid support having a surface to which chemical entities may bind. In some embodiments, guide RNAs being synthesized are attached, directly or indirectly, to the same solid support and may form part of an array. An “array” is a collection of separate molecules of known monomeric sequence each arranged in a spatially defined and a physically addressable manner, such that the location of each sequence is known. An “array,” or “microarray” used interchangeably herein includes any one-dimensional, two-dimensional or substantially two-dimensional (as well as a three-dimensional) arrangement of addressable regions bearing a particular chemical moiety or moieties (such as ligands, e.g., biopolymers such as polynucleotide or oligonucleotide sequences (nucleic acids), polypeptides (e.g., proteins), carbohydrates, lipids, etc.) associated with that region. An array is “addressable” when it has multiple regions of different moieties (e.g., different polynucleotide sequences) such that a region (i.e., a “feature” of the array) at a particular predetermined location (i.e., an “address”) on the array will detect a particular target or class of targets (although a feature may incidentally detect non-targets of that feature). Array features are typically, but need not be, separated by intervening spaces. The number of features that can be contained on an array will largely be determined by the surface area of the substrate, the size of a feature and the spacing between features. Arrays can have densities of up to several hundred thousand or more features per cm², such as 2,500 to 200,000 features/cm². The features may or may not be covalently bonded to the substrate.

Suitable solid supports may have a variety of forms and compositions and derive from naturally occurring materials, naturally occurring materials that have been synthetically modified, or synthetic materials. Examples of suitable support materials include, but are not limited to, silicones, silicon and silicon oxides, teflons, glasses, polysaccharides such as agarose (e.g., Sepharose® from Pharmacia) and dextran (e.g., Sephadex® and Sephacyl®, also from Pharmacia), polyacrylamides, polystyrenes, polyvinyl alcohols, copolymers of hydroxyethyl methacrylate and methyl methacrylate, and the like. In some embodiments, the solid support is a plurality of beads.

The initial monomer of the guide RNAs to be synthesized on the substrate surface can be bound to a linker which in turn is bound to a surface hydrophilic group, e.g., a surface hydroxyl moiety present on a silica substrate. In some embodiments, a universal linker is used. In some other embodiments, the initial monomer is reacted directly with, e.g., a surface hydroxyl moiety. Alternatively, guide RNAs can be synthesized first according to the present invention, and attached to a solid substrate post-synthesis by any

50

method known in the art. Thus, the present invention can be used to prepare arrays of guide RNAs, wherein the oligonucleotides are either synthesized on the array, or attached to the array substrate post-synthesis. Subsequently, the guide RNAs or a pool or a plurality of pools of guide RNAs can optionally and selectively be cleaved from the array substrate and be used as a library or libraries.

IV. Cas Proteins

As mentioned above, a functional CRISPR-Cas system also requires a protein component (e.g., a Cas protein, which may be a Cas nuclease) that provides a desired activity, such as target binding or target nicking/cleaving. In certain embodiments, the desired activity is target binding. In certain embodiments, the desired activity is target nicking or target cleaving. In certain embodiments, the desired activity also includes a function provided by a polypeptide that is covalently fused to a Cas protein, as disclosed herein. In certain embodiments, the desired activity also includes a function provided by a polypeptide that is covalently fused to a nuclease-deficient Cas protein, as disclosed herein. Examples of such a desired activity include a transcription regulation activity (either activation or repression), an epigenetic modification activity, or a target visualization/identification activity, as described below. The Cas protein can be introduced into an in vitro or in vivo system as a purified or non-purified (i) Cas protein or (ii) mRNA encoded for expression of the Cas protein or (iii) linear or circular DNA encoded for expression of the protein. Any of these 3 methods of providing the Cas protein are well known in the art and are implied interchangeably when mention is made herein of a Cas protein or use of a Gas protein. In certain embodiments, the Cas protein is constitutively expressed from mRNA or DNA. In certain embodiments, the expression of Cas protein from mRNA or DNA is inducible or induced.

In certain embodiments, the Cas protein is chemically synthesized (see e.g., Creighton, “Proteins: Structures and Molecular Principles,” W.H. Freeman & Co., NY, 1983), or produced by recombinant DNA technology as described herein. For additional guidance, skilled artisans may consult Frederick M. Ausubel et al., “Current Protocols in Molecular Biology,” John Wiley & Sons, 2003; and Sambrook et al., “Molecular Cloning, A Laboratory Manual,” Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 2001).

In certain embodiments, the Cas protein is provided in purified or isolated form. In certain embodiments, the Cas protein is provided at about 80%, about 90%, about 95%, or about 99% purity. In certain embodiments, the Cas protein is provided as part of a composition. In certain embodiments, the Cas protein is provided in aqueous compositions suitable for use as, or inclusion in, a composition for an RNA-guided nuclease reaction. Those of skill in the art are well aware of the various substances that can be included in such nuclease reaction compositions.

In certain embodiments, a Cas protein is provided as a recombinant polypeptide. In certain examples, the recombinant polypeptide is prepared as a fusion protein. For example, in certain embodiments, a nucleic acid encoding the Cas protein is linked to another nucleic acid encoding a fusion partner, glutathione-s-transferase (GST), 6x-His epitope tag, or M13 Gene 3 protein. Suitable host cells can be used to express the fusion protein. In certain embodiments, the fusion protein is isolated by methods known in the art. In certain embodiments, the fusion protein can be further treated, e.g., by enzymatic digestion, to remove the

US 10,337,001 B2

51

fusion partner and obtain the Cas protein. Alternatively, Cas protein:guide RNA complexes can be made with recombinant technology using a host cell system or an in vitro translation-transcription system known in the art. Details of such systems and technology can be found in e.g., WO2014144761 WO2014144592, WO2013176772, US20140273226, and US20140273233, the contents of which are incorporated herein by reference in their entireties.

Wild Type Cas Proteins

In certain embodiments, a Cas protein comprises a protein derived from a CRISPR-Cas type I, type II, or type III system, which has an RNA-guided polynucleotide binding and/or nuclease activity. Non-limiting examples of suitable Cas proteins include Cas3, Cas4, Cas5, Cas5e (or CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9, Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (or CasA), Cse2 (or CasB), Cse3 (or CasE), Cse4 (or CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Cs17, Cs14, Cs10, Cs16, CsaX, Cs3, Csz1, Cs15, Cf1, Cf2, Cf3, Cf4, and Cu1966. See e.g., WO2014144761 WO2014144592, WO2013176772, US20140273226, and US20140273233, the contents of which are incorporated herein by reference in their entireties.

In certain embodiments, the Cas protein is derived from a type II CRISPR-Cas system. In certain embodiments, the Cas protein is or is derived from a Cas9 protein. In certain embodiments, the Cas protein is or is derived from a bacterial Cas9 protein, including those identified in WO2014144761. In certain embodiments, the Cas protein is or is derived from a *Streptococcus* sp. or *Staphylococcus* sp. Cas9 protein. In certain embodiments, the Cas protein is or is derived from the *Streptococcus thermophilus* Cas9 protein. In certain embodiments, the Cas protein is or is derived from a the *Streptococcus pyogenes* Cas9 protein. In certain embodiments, the Cas protein is or is derived from the *Staphylococcus aureus* Cas9 protein. In certain embodiments, the Cas protein is or is derived from the *Streptococcus thermophilus* Cas9 protein.

In certain embodiments, the wild type Cas protein is a Cas9 protein. In certain embodiments, the wild type Cas9 protein is the Cas9 protein from *S. pyogenes* (SEQ ID NO: 1). In certain embodiments, the protein or polypeptide can comprise, consist of, or consist essentially of a fragment of SEQ ID NO: 1.

In general, a Cas protein includes at least one RNA binding domain, which interacts with the guide RNA. In certain embodiments, the Cas protein is modified to increase nucleic acid binding affinity and/or specificity, alter an enzymatic activity, and/or change another property of the protein. For example, nuclease (i.e., DNase, RNase) domains of the Cas protein can be modified, mutated, deleted, or inactivated. Alternatively, the Cas protein can be truncated to remove domains that are not essential for the function of the protein. In certain embodiments, the Cas protein is truncated or modified to optimize the activity of the effector domain. In certain embodiments, the Cas protein includes a nuclear localization sequence (NLS) that effects importation of the NLS-tagged Cas protein into the nucleus of a living cell. In certain embodiments, the Cas protein includes two or more modifications.

Mutant Cas Proteins

In some embodiments, the Cas protein can be a mutant of a wild type Cas protein (such as Cas9) or a fragment thereof. In other embodiments, the Cas protein can be derived from

52

a mutant Cas protein. For example, the amino acid sequence of the Cas9 protein can be modified to alter one or more properties (e.g., nuclease activity, binding affinity, stability to proteases, etc.) of the protein. Alternatively, domains of the Cas9 protein not involved in RNA-guided cleavage can be eliminated from the protein such that the modified Cas9 protein is smaller than the wild type Cas9 protein. For example, reducing the size of the Cas9 coding sequence can allow it to fit within a transfection vector that otherwise cannot accommodate the wild type sequence, such as the AAV vector among others. In some embodiments, the present system utilizes the Cas protein from *S. pyogenes*, either as encoded in bacteria or codon-optimized for expression in eukaryotic cells. Shown below is the amino acid sequence of wild type *S. pyogenes* Cas9 protein sequence (SEQ ID No. 1, available at www.uniprot.org/uniprot/Q99ZW2).

```

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIG
20 ALLFDSGETAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSF
HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD
KADLRLIYLALAHMIKFRGHFLIEGDLNPNDNSDVDFLFIQLVQTYNQLF
25 EENPINASGVDAKIALSARLSKSRRLENLIAQLPGEKKNGLFGNLIALS
LGLTPNFKNFDLAEDAKLQLSKDTYDDLDNLLAQIGDQYADLFLAAK
NLSDAILLSIDLRLVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQOL
30 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIFKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNRREKIE
KILTFRIPYYVGPLARGNSRFAMTRKSEETITPWNFEVVVDKGASAQS
35 FIERMTNPDKNLNPNEVKLPKHSLLYYEFTVYNELTKVKYVTEGMRKPAF
LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIICKDKFLDNEENEDILEDIVLTLLFEDREMIEERLK
40 TYAHLFDDKVMQLKRRRTGWRGLSRKLINGIRDQSGKTIIDFLKSD
GFANRNFQMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK
GILQTVKVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRI
45 EEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNQGRDMYVDQELDINRL
SDYDVDHIVPQSQLKDDSIDDNKVLTRSDKNRGKSDNVSEEWKFCMKNY
WRQLLNALKITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV
AQILDMSRMTKDYENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINN
50 YHHAHDAYLNAVVGTLLIKYPKLESFVYGDYKVYDVRKMIAKSEOEI
GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR
DFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
55 PKKYGFDSPVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSPEK
NPIDFLEAKGYKEVKKDLIICKLPKYSLFELENGRKRLMASAGELQKGNE
LALPSKYVNFLYLA SHYEKLGKSPEDNEQKQLFVEQHCKHYLDEIIEQIS
60 EFSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAA
FKYFDTTIDRKRYTSTKEVLDATLHQSITGLYETRIDLSQLGGD

```

A Cas9 protein generally has at least two nuclease (e.g., DNase) domains. For example, a Cas9 protein can have a RuvC-like nuclease domain and an HNH-like nuclease domain. The RuvC and HNH domains work together to cut

US 10,337,001 B2

53

both strands in a target site to make a double-stranded break in the target polynucleotide. (Jinek et al., *Science*, 337: 816-821). In certain embodiments, a mutant Cas9 protein is modified to contain only one functional nuclease domain (either a RuvC-like or an HNH-like nuclease domain). For example, in certain embodiments, the mutant Cas9 protein is modified such that one of the nuclease domains is deleted or mutated such that it is no longer functional (i.e., the nuclease activity is absent). In some embodiments where one of the nuclease domains is inactive, the mutant is able to introduce a nick into a double-stranded polynucleotide (such protein is termed a “nickase”) but not able to cleave the double-stranded polynucleotide. For example, an aspartate to alanine (D10A) conversion in a RuvC-like domain converts the Cas9-derived protein into a nickase. Likewise, a histidine to alanine (H840A) conversion in a HNH domain converts the Cas9-derived protein into a nickase. Likewise, an asparagine to alanine (N863A) conversion in a HNH domain converts the Cas9-derived protein into a nickase.

In certain embodiments, both the RuvC-like nuclease domain and the HNH-like nuclease domain are modified or eliminated such that the mutant Cas9 protein is unable to nick or cleave the target polynucleotide. In certain embodiments, all nuclease domains of the Cas9-derived protein are modified or eliminated such that the Cas9-derived protein lacks all nuclease activity. In certain embodiments, a Cas9 protein that lacks some or all nuclease activity relative to a wild-type counterpart, nevertheless, maintains target recognition activity to a greater or lesser extent.

In any of the above-described embodiments, any or all of the nuclease domains can be inactivated by one or more deletion mutations, insertion mutations, and/or substitution mutations using well-known methods, such as site-directed mutagenesis, PCR-mediated mutagenesis, and total gene synthesis, as well as other methods known in the art.

In certain embodiments, the “Cas mutant” or “Cas variant” is at least 50% (e.g., any number between 50% and 100%, inclusive, e.g., 50%, 60%, 70%, 80%, 90%, 95%, 98%, and 99%) identical to SEQ ID NO: 1. In certain embodiments, the “Cas mutant” or “Cas variant” binds to an RNA molecule (e.g., a sgRNA). In certain embodiments, the “Cas mutant” or “Cas variant” is targeted to a specific polynucleotide sequence via the RNA molecule.

Fusion Proteins

In certain embodiments, the Cas protein is fused to another protein or polypeptide heterologous to the Cas protein to create a fusion protein. In certain embodiments, the heterologous sequence includes one or more effector domains, such as a cleavage domain, a transcriptional activation domain, a transcriptional repressor domain, or an epigenetic modification domain. Additional examples of the effector domain include a nuclear localization signal, cell-penetrating or translocation domain, or a marker domain. In certain embodiments, the effector domain is located at the N-terminal, the C-terminal, or in an internal location of the fusion protein. In certain embodiments, the Cas protein of the fusion protein is or is derived from a Cas9 protein. In certain embodiments, the Cas protein of the fusion protein is or is derived from a modified or mutated Cas protein in which all the nuclease domains have been inactivated or deleted. In certain embodiments, the Cas protein of the fusion protein is or is derived from a modified or mutated Cas protein that lacks nuclease activity. In certain embodiments, the RuvC and/or HNH domains of the Cas protein are modified or mutated such that they no longer possess nuclease activity.

54

Cleavage Domains

In certain embodiments, the effector domain of the fusion protein is a cleavage domain. As used herein, a “cleavage domain” refers to a domain that cleaves DNA. The cleavage domain can be obtained from any endonuclease or exonuclease. Non-limiting examples of endonucleases from which a cleavage domain can be derived include restriction endonucleases and homing endonucleases. See, for example, New England Biolabs Catalog or Belfort et al. (1997) *Nucleic Acids Res.* 25, 3379-88. Additional enzymes that cleave DNA are known (e.g., 51 Nuclease; mung bean nuclease; pancreatic DNase I; micrococcal nuclease; yeast HO endonuclease). See also Linn et al. (eds.) “Nucleases,” Cold Spring Harbor Laboratory Press, 1993. One or more of these enzymes (or functional fragments thereof) can be used as a source of cleavage domains.

In certain embodiments, the cleavage domain can be derived from a type II-S endonuclease. Type II-S endonucleases cleave DNA specifically at sites that are typically several base pairs away from the DNA recognition site of the endonuclease and, as such, have separable recognition and cleavage domains. These enzymes generally are monomers that transiently associate to form dimers to cleave each strand of DNA at staggered locations. Non-limiting examples of suitable type II-S endonucleases include BfiI, BpmI, BsAI, BsgI, BsmBI, BsmI, BspMI, FokI, MboII, and SapI. In certain embodiments, the cleavage domain of the fusion protein is a FokI cleavage domain or a fragment or derivative thereof. See Miller et al. (2007) *Nat. Biotechnol.* 25, 778-85; Szepek et al. (2007) *Nat. Biotechnol.* 25, 786-93; Doyon et al. (2011) *Nat. Methods*, 8, 74-81.

Transcriptional Activation Domains

In certain embodiments, the effector domain of the fusion protein is a transcriptional activation domain. In general, a transcriptional activation domain interacts with transcriptional control elements and/or transcriptional regulatory proteins (i.e., transcription factors, RNA polymerases, etc.) to increase and/or activate transcription of a gene. In certain embodiments, the transcriptional activation domain is a herpes simplex virus VP16 activation domain, VP64 (which is a tetrameric derivative of VP16), a NF κ B p65 activation domain, p53 activation domains 1 and 2, a CREB (cAMP response element binding protein) activation domain, an E2A activation domain, or an NFAT (nuclear factor of activated T-cells) activation domain. In certain embodiments, the transcriptional activation domain is Gal4, Gen4, MLL, Rtg3, Gln3, Oaf1, Pip2, Pdr1, Pdr3, Pho4, or Leu3. The transcriptional activation domain may be wild type, or it may be a modified or truncated version of the original transcriptional activation domain.

Transcriptional Repressor Domains

In certain embodiments, the effector domain of the fusion protein is a transcriptional repressor domain. In general, a transcriptional repressor domain interacts with transcriptional control elements and/or transcriptional regulatory proteins (i.e., transcription factors, RNA polymerases, etc.) to decrease and/or prohibit transcription of a gene. In certain embodiments, the transcriptional repressor domains is inducible cAMP early repressor (ICER) domains, Kruppel-associated box A (KRAB-A) repressor domains, YY1 glycine rich repressor domains, Sp1-like repressors, E(spl) repressors, IkB repressor, or MeCP2.

Epigenetic Modification Domains

In certain embodiments, the effector domain of the fusion protein is an epigenetic modification domain. In general, epigenetic modification domains alter gene expression by modifying the histone structure and/or chromosomal struc-

US 10,337,001 B2

55

ture. In certain embodiments, the epigenetic modification domains is a histone acetyl transferase domain, a histone deacetylase domain, a histone methyltransferase domain, a histone demethylase domain, a DNA methyltransferase domain, or a DNA demethylase domain.

Additional Domains

In certain embodiments, the fusion protein further comprises at least one additional domain. Non-limiting examples of suitable additional domains include nuclear localization signals (NLSs), cell-penetrating or translocation domains, and marker domains. An NLS generally comprises a stretch of basic amino acids. See, e.g., Lange et al. (2007) *J. Biol. Chem.*, 282, 5101-5. For example, in certain embodiments, the NLS is a monopartite sequence, such as PKKKRKV (SEQ ID NO: 2) or PKKRRV (SEQ ID NO: 3). In certain embodiments, the NLS is a bipartite sequence. In certain embodiments, the NLS is KRPAATKKAGQAKKK (SEQ ID NO: 4).

In certain embodiments, the fusion protein comprises at least one cell-penetrating domain. In certain embodiments, the cell-penetrating domain is a cell-penetrating peptide sequence derived from the HIV-1 TAT protein. As an example, the TAT cell-penetrating sequence can be GRK-KRRQRRRPPQPKKKRKV (SEQ ID NO: 5). In certain embodiments, the cell-penetrating domain is TLM (PLSS-BFSRIDDPPKKKRKV; SEQ ID NO: 6), a cell-penetrating peptide sequence derived from the human hepatitis B virus. In certain embodiments, the cell-penetrating domain is MPG (GALFLGWLGAAGSTMGAPKKRKV; SEQ ID NO: 7 or GALFLGFLGAAGSTMGAWSQPKKKRKV; SEQ ID NO: 8). In certain embodiments, the cell-penetrating domain is Pcp-1 (KETWWETWWTEWSQPKKKRKV; SEQ ID NO: 9), VP22, a cell penetrating peptide from Herpes simplex virus, or a polyarginine peptide sequence.

In certain embodiments, the fusion protein comprises at least one marker domain. Non-limiting examples of marker domains include fluorescent proteins, purification tags, and epitope tags. In certain embodiments, the marker domain is a fluorescent protein. Non limiting examples of suitable fluorescent proteins include green fluorescent proteins (e.g., GFP, GFP-2, tagGFP, turbGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), yellow fluorescent proteins (e.g., YFP, EYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), blue fluorescent proteins (e.g. EBFP, EBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-sapphire,), cyan fluorescent proteins (e.g. ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), red fluorescent proteins (mKate, mKate2, mPlum, DsRed monomer, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRc1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), orange fluorescent proteins (mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato) and any other suitable fluorescent protein. In certain embodiments, the marker domain is a purification tag and/or an epitope tag. Exemplary tags include, but are not limited to, glutathione-S-transferase (GST), chitin binding protein (CBP), maltose binding protein, thioredoxin (TRX), poly(NANP), tandem affinity purification (TAP) tag, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, biotin carboxyl carrier protein (BCCP), and calmodulin.

V. Uses and Methods

In one aspect, the present invention provides a method for cleaving a target polynucleotide with a Cas protein. The

56

method comprises contacting the target polynucleotide with (i) a guide RNA or a set of guide RNA molecules described herein, and (ii) a Cas protein. In certain embodiments, the method results in a double-strand break in the target polynucleotide. In certain embodiments, the Cas protein is a Cas protein having a single-strand nicking activity. In certain embodiments, the method results in a single-strand break in the target polynucleotide. In certain embodiments, a complex comprising a guide RNA and Cas protein having a 10 single-strand nicking activity is used for sequence-targeted single-stranded DNA cleavage, i.e., nicking.

In one aspect, the present invention provides a method for cleaving two or more target polynucleotides with a Cas protein. The method comprises contacting the target polynucleotides with (i) a set of guide RNA molecules described herein, and (ii) a Cas protein. In certain embodiments, the method results in double-strand breaks in the target polynucleotides. In certain embodiments, the Cas protein is a Cas protein having a single-strand nicking activity. In certain 15 embodiments, the method results in single-strand breaks in the target polynucleotides. In certain embodiments, a complex comprising a guide RNA and Cas protein having a single-strand nicking activity is used for sequence-targeted single-stranded DNA cleavage, i.e., nicking.

In one aspect, the present invention provides a method for binding a target polynucleotide with a Cas protein. The method comprises contacting the target polynucleotide with (i) a guide RNA or a set of guide RNA molecules described herein and (ii) a Cas protein, to result in binding of the target polynucleotide with the Cas protein. In certain embodiments, the Cas protein is a Cas variant. In certain embodiments, the Cas variant lacks some or all nuclease activity relative to a counterpart wild-type Cas protein.

In one aspect, the present invention provides a method for 35 binding two or more target polynucleotides with a Cas protein. The method comprises contacting the target polynucleotides with (i) a set of RNA molecules described herein and (ii) a Cas protein, to result in binding of the target polynucleotides with the Cas protein. In certain 40 embodiments, the Cas protein is a Cas variant. In certain embodiments, the Cas variant lacks some or all nuclease activity relative to a counterpart wild-type Cas protein.

In one aspect, the present invention provides a method for targeting a Cas protein to a target polynucleotide. The 45 method comprises contacting the Cas protein with a guide RNA or a set of guide RNA molecules described herein. In certain embodiments, the method results in formation of a guide RNA:Cas protein complex. In certain embodiments, the Cas protein is a wild type Cas9 protein. In certain embodiments, the Cas protein is a mutant or variant of a Cas9 protein. In certain embodiments, the Cas protein is a Cas protein having a single-strand nicking activity. In certain 50 embodiments, the Cas protein is a Cas protein lacking nuclease activity (e.g., a nuclease-deficient mutant of Cas protein). In certain embodiments, the Cas protein is part of a fusion protein (e.g., a fusion protein comprising (i) the Cas protein and (ii) a heterologous polypeptide).

In one aspect, the present invention provides a method for targeting a Cas protein to two or more target polynucleotides. The method comprises contacting the Cas protein with a set of guide RNA molecules described herein. In certain 55 embodiments, the method results in formation of a guide RNA:Cas protein complex. In certain embodiments, the Cas protein is a wild type Cas9 protein. In certain embodiments, the Cas protein is a mutant or variant of a Cas9 protein. In certain embodiments, the Cas protein is a Cas protein having a single-strand nicking activity. In certain

US 10,337,001 B2

57

embodiments, the Cas protein is a Cas protein lacking nuclease activity (e.g., a nuclease-deficient mutant of Cas protein). In certain embodiments, the Cas protein is part of a fusion protein (e.g., a fusion protein comprising (i) the Cas protein or and (ii) a heterologous polypeptide).

In certain embodiments, the guide RNA is introduced into a cell by transfection. Techniques for RNA transfection are known in the art and include electroporation and lipofection. Effective techniques for RNA transfection depend mostly on cell type. See, e.g., Lujambio et al. (Spanish National Cancer Centre) *Cancer Res.* February 2007, which describes transfection of HTC-116 colon cancer cells and uses Oligofectamine (Invitrogen) for transfection of commercially obtained, modified miRNA or precursor miRNA. See also, Cho et al. (Seoul National Univ.) *Nat. Biotechnol.* March 2013, which describes transfection of K562 cells and uses 4D Nucleofection™ (Lonza) electroporation for transfection of transcribed sgRNAs (about 60 nts long). Techniques for transfection of RNA are also known in the art. For example, therapeutic RNA has been delivered in non-pathogenic *E. coli* coated with Invasin protein (to facilitate uptake into cells expressing β-1 integrin protein) and with the *E. coli* encoded to express lysteriolysin O pore-forming protein to permit the shRNA to pass from the *E. coli* into the cytoplasm. See also Cho et al. (Seoul National Univ.) *Nat. Biotechnol.* March 2013.

In certain embodiments, the guide RNA is introduced or delivered into cells. Technologies that can be used for delivery of guide RNA include those that utilize encapsulation by biodegradable polymers, liposomes, or nanoparticles. Such polymers, liposomes, and nanoparticles can be delivered intravenously. In certain embodiments, for in vivo delivery, guide RNA can be injected into a tissue site or administered systemically. In vivo delivery can also be effected by a beta-glucan delivery system, such as those described in U.S. Pat. Nos. 5,032,401 and 5,607,677, and U.S. Publication No. 2005/0281781, which are hereby incorporated by reference in their entirety. In certain embodiments, guide RNA or a delivery vehicle containing guide RNA is targeted to a particular tissue or body compartment. For example, in certain embodiments, to target exogenous RNA to other tissues, synthetic carriers are decorated with cell-specific ligands or aptamers for receptor uptake, e.g., RNA encased in cyclodextrin nanoparticles coated with PEG and functionalized with human transferrin protein for uptake via the transferrin receptor which is highly expressed in tumor cells. Further approaches are described herein below or known in the art.

The present invention has been tested in human cells as described in Hendel et al., *Nat. Biotechnol.* (2015) 33:9, 985-9 (which is incorporated in this application in its entirety). In the cited work, modified guide RNA was introduced into K562 cells, human primary T cells, and CD34+ hematopoietic stem and progenitor cells (HSPCs). The modified guide RNA significantly enhanced genome editing efficiencies in human cells, including human primary T cells and CD34+ HSPCs as compared to unmodified guide RNA.

FIGS. 11A and 11B illustrate experimental results showing that gene disruption in human cell lines can be achieved by high frequencies of indels or by cleavage-stimulated homologous recombination using synthesized and chemically modified sgRNAs disclosed herein. Gene disruption by mutagenic NHEJ was measured by deep sequencing of PGR amplicons (FIG. 12A) or gene addition by HR at the three loci IL2RG, HBB and CCR5 in K562 cells induced by Cas9 in combination with, synthetic sgRNAs (FIG. 12B). The

58

synthetic sgRNAs were delivered at 1 µg (light shade) or 20 µg (dark shade) per 1 million cells. Cas9 was expressed from a plasmid (2 µg) and for HR experiments 5 µg of GFP-encoding donor plasmid was included. As a positive control, 5 2 µg of sgRNA plasmid encoding both the sgRNA and the Cas9 protein was used (gray bars). Bars represent average values+s.e.m., n=3.

FIGS. 12A, 12B, 12C and 12D illustrate experimental results showing that chemically modified sgRNAs as 10 described herein can be used to achieve high frequencies of gene disruption or targeted genome editing in stimulated primary human T cells and CD34+ hematopoietic stem and progenitor cells (HSPCs).

FIG. 12A illustrates results from primary human T cells 15 nucleofected with 10 µg of a synthetic CCR5 sgRNAs and either 15 µg Cas9 mRNA or 1 µg Cas9-encoding plasmid. 1 µg sgRNA plasmid encoding both the sgRNA and Cas9 protein was included for comparison. The bars represent average indel frequencies for three different donors+s.e.m., 20 n=6, as measured by TIDE (tracking of indels by decomposition) analysis of PCR amplicons spanning the sgRNA target site, and using a mock-treated sample as control reference. Delivery of Cas9 mRNA with the unmodified or the M-modified sgRNA, and nucleofection of the plasmid 25 encoding both the sgRNA and Cas9, did not give rise to allele modification frequencies above background. Co-transfection of the MSP-modified sgRNA with DNA expression plasmid for Cas9 generated 9.3% indel frequency. Cas9 mRNA with either the MS- or MSP-modified sgRNA generated 30 48.7% and 47.9% indel frequencies, respectively.

FIG. 12B illustrates results from stimulated T cells. The 35 cells were nucleofected as above, but with 15 µg Cas9 protein complexed with a 2.5 molar excess of the indicated synthetic CCR5 sgRNAs. Indel frequencies were measured 40 by TIDE analysis. The bars represent average indel frequencies for three different donors+s.e.m., n=6. A 2.4-fold improvement in indel frequencies of the MS-modified sgRNA over the unmodified sgRNA (30.7% vs. 12.8%) was observed for chemically modified sgRNAs when delivered 45 complexed with Cas9 protein. These results establish that chemically modified sgRNAs can be used for genome editing of stimulated T cells when delivered complexed with Cas9 protein.

FIG. 12C illustrates results from human peripheral blood 50 CD34+ HSPCs. 500,000 mobilized cells were nucleofected with 10 µg of the indicated synthetic sgRNAs targeting IL2RG or HBB and either 15 µg Cas9 mRNA or 1 µg Cas9 plasmid. 1 µg of sgRNA plasmid encoding both the sgRNA and Cas9 protein was included for comparison. Bars represent average indel frequencies+s.e.m., n=3, as measured by 55 T7 endonuclease cleavage assay. Indels were not detected at either locus using the unmodified or M-modified sgRNAs when co-transfected with Cas9 mRNA. However, the IL2RG MS- and MSP-modified sgRNAs showed 17.5% and 17.7% indel frequencies, respectively, and 23.4% and 22.0%, respectively, for the HBB MS- and MSP-modified sgRNAs.

FIG. 12D illustrates results from stimulated T cells or 60 mobilized human peripheral blood CD34+ HSPCs. One million cells were nucleofected with 15 µg Cas9 mRNA and 10 µg of the indicated synthetic CCR5 sgRNAs. A recent study showed that the simultaneous use of two sgRNAs could improve gene disruption in human primary T cells and in CD34+ HSPCs. See, e.g., Mandal et al. (2014) *Cell Stem Cell*, 15, 643-52. MS- and MSP-modified CCR5 sgRNAs 65 were chemically synthesized with the sequences reported in Mandal study (termed 'D' and 'Q'), which cut 205 base pairs

US 10,337,001 B2

59

apart. When used in combination, the amount of each sgRNA was 5 µg. Indel frequencies for samples with single sgRNAs were measured by TIDE analysis as above and allele disruption frequencies for samples with two sgRNAs were measured by sequencing of cloned PCR products. The bars represent average indel frequencies+s.e.m., n=3. In T cells, the 'D' sgRNA alone gave rise to 56.0% and 56.3% indels for the MS- and MSP-modified sgRNA, respectively, and the 'Q' sgRNA gave rise to 62.6% and 69.6% indels, respectively. When used in combination, the frequencies of allele modification increased, as we observed 73.9% and 93.1% indels for the MS- and MSP-modified sgRNAs, respectively, of which the majority of the modification events were deletions between the two sgRNA target sites. In CD34+ HSPCs, observations were similar though the overall frequencies were lower. For the 'D' sgRNA, allele modification frequencies of 9.8% and 11.2% were observed for the MS- and MSP-modified sgRNA, respectively, and 17.8% and 19.2% for the 'Q' sgRNA. When used in combination the frequencies increased to 37.8% and 43.0% for the MS- and MSP-modified sgRNAs, respectively. This shows that the use of two chemically modified sgRNAs is a highly effective way to facilitate gene disruption in primary human T cells and CD34+ HSPCs.

Examples of other uses include genomic editing and gene expression regulation as described below.

Genomic Editing

In one aspect, the present invention provides a method for genomic editing to modify a DNA sequence in vivo or in vitro ("in vitro" includes, without being limited to, a cell-free system, a cell lysate, an isolated component of a cell, and a cell outside of a living organism). The DNA sequence may comprise a chromosomal sequence, an episomal sequence, a plasmid, a mitochondrial DNA sequence, or a functional intergenic sequence, such, as an enhancer sequence or a DNA sequence for a non-coding RNA. The method comprises contacting the DNA sequence with (i) a guide RNA or a set of guide RNA molecules described herein, and (ii) a Cas protein. In certain embodiments, the DNA sequence is contacted outside of a cell. In certain embodiments, the DNA sequence is located in the genome within a cell and is contacted in vitro or in vivo. In certain embodiments, the cell is within an organism or tissue. In certain embodiments, the cell is a human cell, a non-human mammalian cell, a stem cell, a non-mammalian vertebrate cell, an invertebrate cell, a plant cell, a single cell organism, or an embryo. In certain embodiments, the guide RNA aids in targeting the Cas protein to a targeted site in the DNA sequence. In certain embodiments, the Cas protein cleaves at least one strand of the DNA sequence at the targeted site. In certain embodiments, the Cas protein cleaves both strands of the DNA sequence at the targeted site.

In certain embodiments, the method further comprises introducing the Cas protein into a cell or another system. In certain embodiments, the Cas protein is introduced as a purified or non-purified protein. In certain embodiments, the Cas protein is introduced via an mRNA encoding the Cas protein. In certain embodiments, the Cas protein is introduced via a linear or circular DNA encoding the Cas protein. In certain embodiments, the cell or system comprises a Cas protein or a nucleic acid encoding a Cas protein.

In certain embodiments, a double-stranded break can be repaired via an error-prone, non-homologous end-joining ("NHEJ") repair process. In certain embodiments, a double-stranded break can be repaired by a homology-directed

60

repair (HDR) process such that a donor sequence in a donor polynucleotide can be integrated into or exchanged with the targeted DNA sequence.

In certain embodiments, the method further comprises introducing at least one donor polynucleotide into the cell or system. In certain embodiments, the donor polynucleotide comprises at least one homologous sequence having substantial sequence identity with a sequence, on either side of the targeted site in the DNA sequence. In certain embodiments, the donor polynucleotide comprises a donor sequence that can be integrated into or exchanged with the DNA sequence via homology-directed repair, such as homologous recombination.

In certain embodiments, the donor polynucleotide includes an upstream homologous sequence and a downstream homologous sequence, each of which have substantial sequence identity to sequences located upstream and downstream, respectively, of the targeted site in the DNA sequence. These sequence similarities permit, for example, homologous recombination between the donor polynucleotide and the targeted DNA sequence such that the donor sequence can be integrated into (or exchanged with) the DNA sequence targeted.

In certain embodiments, the target site(s) in the DNA sequence spans or is adjacent to a mutation, e.g., point mutation, a translocation or an inversion which may cause or be associated with a disorder. In certain embodiments, the method comprises correcting the mutation by introducing into the cell or system at least one donor polynucleotide comprising (i) a wild type counterpart of the mutation and (ii) at least one homologous sequence having substantial sequence identity with a sequence on one side of the targeted site in the DNA sequence. In certain embodiments, the donor polynucleotide comprises a homologous sequence having substantial sequence identity with a sequence on both sides of the targeted site in the DNA sequence.

In certain embodiments, the donor polynucleotide comprises an exogenous sequence that can be integrated into or exchanged with the targeted DNA sequence via a homology-directed repair process, such as homologous recombination. In certain embodiments, the exogenous sequence comprises a protein coding gene, which, optionally, is operably linked to an exogenous promoter control sequence. Thus, in certain embodiments, upon integration of the exogenous sequence, a cell can express a protein encoded by the integrated gene. In certain embodiments, the exogenous sequence is integrated into the targeted DNA sequence such that its expression in the recipient cell or system is regulated by the exogenous promoter control sequence. Integration of an exogenous gene into the targeted DNA sequence is termed a "knock in." In other embodiments, the exogenous sequence can be a transcriptional control sequence, another expression control sequence, an RNA coding sequence, and the like.

In certain embodiments, the donor polynucleotide comprises a sequence that is essentially identical to a portion of the DNA sequence at or near the targeted site, but comprises at least one nucleotide change. For example, in certain embodiments, the donor sequence comprises a modified or mutated version of the DNA sequence at or near the targeted site such that, upon integration or exchange with the targeted site, the resulting sequence at the targeted site comprises at least one nucleotide change. In certain embodiments, the at least one nucleotide change is an insertion of one or more nucleotides, a deletion of one or more nucleotides, a substitution of one or more nucleotides, or combinations thereof. As a consequence of the integration of the modified

US 10,337,001 B2

61

sequence, the cell may produce a modified gene product from the targeted DNA sequence.

In certain embodiments, the methods are for multiplex applications. In certain embodiments, the methods comprise introducing a library of guide RNAs into the cell or system. In certain embodiments, the library comprises at least 100 unique guide sequences. In certain embodiments, the library comprises at least 1,000 unique guide sequences. In certain embodiments, the library comprises at least 10,000 unique guide sequences. In certain embodiments, the library comprises at least 100,000 unique guide sequences. In certain embodiments, the library comprises at least 1,000,000 unique guide sequences. In certain embodiments, the library targets at least 10 different polynucleotides or at least 10 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 100 different polynucleotides or at least 100 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 1,000 different polynucleotides or at least 1,000 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 10,000 different polynucleotides or at least 10,000 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 100,000 different polynucleotides or at least 100,000 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 1,000,000 different polynucleotides or at least 1,000,000 different sequences within one or more polynucleotides.

Genomic Editing in Human and Mammalian Cells

Embodiments of the present invention are useful in methods for genomic editing to modify a target polynucleotide, for example a DNA sequence, in a mammalian cell.

In certain embodiments, the DNA sequence is a chromosomal sequence. In certain embodiments, the DNA sequence is a protein-coding sequence. In certain embodiments, the DNA sequence is a functional intergenic sequence, such as an enhancer sequence or a non-coding sequence. In certain embodiments, the DNA is part of a human gene. In some such embodiments, the human gene is the clathrin light chain (CLTA1) gene, the human interleukin 2 receptor gamma (IL2RG) gene, the human cytotoxic T-lymphocyte-associated protein 4 (CLTA4) gene, the human protocadherin alpha 4 (PCDHA4) gene, the human engrailed homeobox 1 (EN1) gene, the human hemoglobin beta (HBB) gene, which can harbor mutations responsible for sickle cell anemia and thalassemias, or the human chemokine (C-C motif) receptor 5 (CCR5) gene which encodes a co-receptor of HIV.

In certain embodiments, the mammalian cell is a human cell. In some such embodiments, the human cell is a primary human cell. In further embodiments, the primary human cell is a human primary T cell. The human primary T cell may be stimulated or unstimulated. In certain embodiments, the human cell is a stem/progenitor cell, such as a CD34+ hematopoietic stem and progenitor cell (HSPC). In certain embodiments, the human cell is from a cultured cell line, for example such as can be obtained commercially. Exemplary cell lines include K562 cells, a human myelogenous leukemia line.

In certain embodiments, the cell is within a living organism. In certain other embodiments, the cell is outside of a living organism.

The method comprises contacting the DNA sequence with (i) a guide RNA or a set of guide RNA molecules described herein, and (ii) a Cas protein.

62

In certain embodiments, the method further comprises introducing or delivering the guide RNA into the cell. In some such embodiments, the guide RNA is introduced into a cell by transfection. Techniques for RNA transfection are known in the art and include electroporation and lipofection. In other embodiments, the guide RNA is introduced into a cell (and, more particularly, a cell nucleus) by nucleofection. Techniques for nucleofection are known in the art and may utilize nucleofection devices such as the Lonza Nucleofector 2b or the Lonza 4D-Nucleofector and associated reagents.

In certain embodiments, the method further comprises introducing or delivering the Cas protein into the cell. In some such embodiments, the Cas protein is introduced as a purified or non-purified protein. In other embodiments, the Cas protein is introduced via an mRNA encoding the Cas protein. In some such embodiments, the mRNA encoding the Cas protein is introduced into the cell by transfection. In other embodiments, the mRNA encoding the Cas protein is introduced into a cell (and, more particularly, a cell nucleus) by nucleofection.

In certain embodiments, the method employs ribonucleoprotein (RNP)-based delivery such that the Cas protein is introduced into the cell in a complex with the guide RNA. For example, a Cas9 protein may be complexed with a guide RNA in a Cas9:gRNA complex, which allows for co-delivery of the gRNA and Cas protein. For example, the Cas:gRNA complex may be nucleofected into cells.

In certain embodiments, the method employs an all-RNA delivery platform. For example, in some such embodiments, the guide RNA and the mRNA encoding the Cas protein are introduced into the cell simultaneously or substantially simultaneously (e.g., by co-transfection or co-nucleofection). In certain embodiments, co-delivery of Cas mRNA and modified gRNA results in higher editing frequencies as compared to co-delivery of Cas mRNA and unmodified gRNA. In particular, gRNA having 2'-O-methyl-3'-phosphorothioate (MS), or 2'-O-methyl-3'-thioPACE (MSP) incorporated at three terminal nucleotides at both the 5' and 3' ends, provide higher editing frequencies as compared to unmodified gRNA.

In certain embodiments, the guide RNA and the mRNA encoding the Cas protein are introduced into the cell sequentially; that is, the guide RNA and the mRNA encoding the Cas protein are introduced into the cell at different times. The time period between the introduction of each agent may range from a few minutes (or less) to several hours or days. For example, in some such embodiments, gRNA is delivered first, followed by delivery of Cas mRNA 4, 8, 12 or 24 hours later. In other such embodiments, Cas mRNA is delivered first, followed by delivery of gRNA 4, 8, 12 or 24 hours later. In some particular embodiments, delivery of modified gRNA first, followed by delivery of Cas mRNA results in Higher editing frequencies as compared to delivery of unmodified gRNA followed by delivery of Cas mRNA.

In certain embodiments, the gRNA is introduced into the cell together with a DNA plasmid encoding the Cas protein. In some such embodiments, the gRNA and the DNA plasmid encoding the Cas protein are introduced into the cell by nucleofection. In some particular embodiments, an RNP-based delivery platform or an all-RNA delivery platform provides lower cytotoxicity in primary cells than a DNA plasmid-based delivery system.

In certain embodiments, the method provides significantly enhanced genome editing efficiencies in human cells, including human primary T cells and CD34+ HSPCs.

In certain embodiments, modified gRNA increases the frequency of insertions or deletions (indels), which may be

US 10,337,001 B2

63

indicative of mutagenic NHEJ and gene disruption, relative to unmodified gRNA. In particular, modified gRNA having 2'-O-methyl-3'-phosphorothioate (MS) or 2'-O-methyl-3'-thioPACE (MSP) incorporated at three terminal nucleotides at both the 5' and 3' ends, increases the frequency of indels relative to unmodified gRNA.

In certain embodiments, co-delivery of modified gRNA and Cas mRNA to human primary T cells increases the frequency of indels as compared to co-delivery of unmodified gRNA and Cas mRNA. In particular, modified gRNA having 2'-O-methyl-3'-phosphorothioate (MS) or 2'-O-methyl-3'-thioPACE (MSP) incorporated at three terminal nucleotides at both the 5' and 3' ends, increases the frequency of indels in human primary T cells relative to unmodified gRNA.

In certain embodiments, modified gRNA improves gRNA stability relative to unmodified gRNA. As one example, gRNA having 2'-O-methyl (M) incorporated at three terminal nucleotides at both the 5' and 3' ends, modestly improves stability against nucleases and also improves base pairing thermostability over unmodified gRNA. As another example, gRNA having 2'-O-methyl-3'-phosphorothioate (MS) or 2'-O-methyl-3'-thioPACE (MSP) incorporated at three terminal nucleotides at both the 5' and 3' ends, dramatically improves stability against nucleases relative to unmodified gRNA. It is contemplated that gRNA end modifications enhance intracellular stability against exonucleases, thus enabling increased efficacy of genome editing when Cas mRNA and gRNA are co-delivered or sequentially delivered into human cells.

In certain embodiments, modified gRNA stimulates gene targeting, which, in turn, allows for gene editing by, for example, homologous recombination or NHEJ. In particular, gRNA having 2'-O-methyl-3'-phosphorothioate (MS), or 2'-O-methyl-3'-thioPACE (MSP) incorporated at three terminal nucleotides at both the 5' and 3' ends, stimulates higher levels of homologous recombination than unmodified gRNA.

In certain embodiments, modified gRNA retains high specificity. In certain embodiments, the ratio of on-target to off-target indel frequencies is improved with modified gRNA as compared to unmodified gRNA. In certain embodiments, modified gRNA delivered in an RNP complex with a Cas protein provides significantly better on-target: off-target ratios compared to a DNA plasmid-based delivery system.

Gene Expression Regulation

In certain embodiments, the guide RNA described herein is used for regulating transcription or expression of a gene of interest. For example, in certain embodiments, a fusion protein comprising a Cas protein (e.g., a nuclease-deficient Cas9) and a transcription activator polypeptide is used to increase transcription of a gene. Similarly, in certain embodiments, a fusion protein comprising a Cas protein (e.g., a nuclease-deficient Cas9) and a repressor polypeptide is used to knock-down gene expression by interfering with transcription of the gene.

In at least one aspect, the present invention provides a method for regulating the expression of a gene of interest in vivo or in vitro. The method comprises introducing into a cell or another system (i) a synthetic guide RNA described herein, and (ii) a fusion protein. In certain embodiments, the fusion protein comprises a Cas protein and an effector domain, such as a transcriptional activation domain, a transcriptional repressor domain, or an epigenetic modification domain. In certain embodiments, the fusion protein comprises a mutated Cas protein, such as a Cas9 protein that is

64

a null nuclease. In certain embodiments, the Cas protein contains one or more mutations, such as D10A, H840A and/or N863A.

In certain embodiments, the fusion protein is introduced into the cell or system as a purified or non-purified protein. In certain embodiments, the fusion protein is introduced into the cell or system via an mRNA encoding the fusion protein. In certain embodiments, the fusion protein is introduced into the cell or system via a linear or circular DNA encoding the fusion protein.

In certain embodiments, the guide RNA aids in directing the fusion protein to a specific target polynucleotide comprising a chromosomal sequence, an episomal sequence, a plasmid, a mitochondrial DNA sequence, or a functional intergenic sequence, such as an enhancer or the DNA sequence for a non-coding RNA. In certain embodiments, the effector domain regulates expression of a sequence in the target polynucleotide. A guide RNA for modulating gene expression can be designed to target any desired endogenous gene or sequence encoding a functional RNA. A genomic target sequence can be selected in proximity of the transcription start site of the endogenous gene, or alternatively, in proximity of the translation initiation site of the endogenous gene. In certain embodiments, the target sequence is in a region of the DNA that is traditionally termed the promoter proximal region of a gene. In certain embodiments, the target sequence lies in a region from about 1,000 base pairs upstream of the transcription start site to about 1,000 base pairs downstream of the transcription start site. In certain embodiments, the target sequence is remote from the start site for transcription of the gene (e.g., on another chromosome).

In certain embodiments, the methods are for multiplex applications. In certain embodiments, the methods comprise introducing a library of guide RNAs into the cell or system. In certain embodiments, the library comprises at least 100, at least 1,000, at least 10,000, at least 100,000, or at least 1,000,000 unique guide sequences. In certain embodiments, the library targets at least 10 different polynucleotides or at least 10 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 100 different polynucleotides or at least 100 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 1,000 different polynucleotides or at least 1,000 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 10,000 different polynucleotides or at least 10,000 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 100,000 different polynucleotides or at least 100,000 different sequences within one or more polynucleotides. In certain embodiments, the library targets at least 1,000,000 different polynucleotides or at least 1,000,000 different sequences within one or more polynucleotides.

Kits

In one aspect, the present invention provides kits containing reagents for performing the above-described methods, including producing gRNA:Cas protein complex and/or supporting its activity for binding, nicking or cleaving target polynucleotide. In certain embodiments, one or more of the reaction components, e.g., one or more guide RNAs and Cas proteins, for the methods disclosed herein, can be supplied in the form of a kit for use. In certain embodiments, the kit comprises a Cas protein or a nucleic acid encoding the Cas protein, and one or more guide RNAs described herein or a set or library of guide RNAs. In certain embodiments, the kit includes one or more other reaction components. In certain

US 10,337,001 B2

65

embodiments, an appropriate amount of one or more reaction components is provided in one or more containers or held on a substrate.

Examples of additional components of the kits include, but are not limited to, one or more different polymerases, one or more host cells, one or more reagents for introducing foreign nucleic acid into host cells, one or more reagents (e.g., probes or PCR primers) for detecting expression of the guide RNA and/or the Cas mRNA or protein or for verifying the status of the target nucleic acid, and buffers, transfection reagents or culture media for the reactions (in 1× or more concentrated forms). In certain embodiments, the kit includes one or more of the following components: biochemical and physical supports; terminating, modifying and/or digesting reagents; osmolytes; and apparatus for reaction, transfection and/or detection.

The reaction components used can be provided in a variety of forms. For example, the components (e.g., enzymes, RNAs, probes and/or primers) can be suspended in an aqueous solution or bound to a bead or as a freeze-dried or lyophilized powder or pellet. In the latter case, the components, when reconstituted, form a complete mixture of components for use in an assay. The kits of the invention can be provided at any suitable temperature. For example, for storage of kits containing protein components or complexes thereof in a liquid, it is preferred that they are provided and maintained below 0° C., preferably at about -20° C., possibly in a freeze-resistant solution containing glycerol or other suitable antifreeze.

A kit or system may contain, in an amount sufficient for at least one assay, any combination of the components described herein. In some applications, one or more reaction components may be provided in pre-measured single use amounts in individual, typically disposable, tubes or equivalent containers. With such an arrangement, a RNA-guided nuclease reaction can be performed by adding a target nucleic acid, or a sample or cell containing the target nucleic acid, to the individual tubes directly. The amount of a component supplied in the kit can be any appropriate amount and may depend on the market to which the product is directed. The container(s) in which the components are supplied can be any conventional container that is capable of holding the supplied form, for instance, microfuge tubes, microtiter plates, ampoules, bottles, or integral testing devices, such as fluidic devices, cartridges, lateral flow, or other similar devices.

The kits can also include packaging materials for holding the container or combination of containers. Typical packaging materials for such kits and systems include solid matrices (e.g., glass, plastic, paper, foil, micro-particles and the like) that hold the reaction components or detection probes in any of a variety of configurations (e.g., in a vial, microtiter plate well, microarray, and the like). The kits may further include instructions recorded in a tangible form for use of the components.

EXAMPLES

Example 1

To evaluate the ability of the chemically synthesized guide RNAs to target and cleave a DNA target sequence, an in vitro cleavage assay was developed. Briefly, as shown in FIG. 3, ~4-kb PAM-addressable DNA targets were prepared by preparative PGR amplification of plasmid-borne human sequences (here, a sequence from the human clathrin light chain CLTA gene). In a 20-uL reaction volume, 50 fmoles of

66

linearized DNA target in the presence of 50 nM sgRNA, 39 nM recombinant purified Cas9 protein (*S. pyogenes*; Agilent) and 10 mM MgCl₂ at pH 7.6 was incubated at 37° C. for 30 min. Upon completion, 0.5 uL of RNase It (Agilent) was added, and incubation was continued at 37° C. for 5 min and then at 70° C. for 15 min. Subsequently 0.5 uL of Proteinase K (Mol. Bio. grade, NEB) was added and incubated at 37° C. for 15 min. Aliquots were loaded into a DNA 7500 LabChip and were analyzed on a Bioanalyzer 2200.

10 The workup steps served to release Cas9 from binding to target DNA, which were assayed for cleavage.

A series of guide RNAs as listed in FIG. 4 were chemically synthesized. Briefly, individual RNA strands were synthesized and HPLC purified. All oligonucleotides were 15 quality control approved on the basis of chemical purity by HPLC analysis and full-length strand purity by mass spectrometry analysis. Each of these guide RNAs was designed to target the human CLTA gene.

The results are shown in FIG. 4. As shown in the Table 1 20 of FIG. 4, all but one of the chemically synthesized guide RNAs targeted and cleaved the CLTA-en coded DNA target sequence with significant cleavage rates. The one exception was "CLTA_37_Deoxy" guide RNA, which had a contiguous 25 sequence of 37 deoxyribonucleotides at its 5' end.

As disclosed herein, a variety of chemical modifications 30 were tested at specific positions in the sequence of a guide RNA. Surprisingly, the tested positions in the guide sequence of the guide RNA (a.k.a. the spacer sequence in the guide RNA) tolerated most of the modifications tested, including combinations of multiple modifications within single nucleotides in the guide RNA, even when modifications 35 were instantiated in the target-binding sequences.

The results revealed that guide RNAs containing modifications at specific positions were tolerated by active Cas 35 protein and gRNA: Cas protein complexes, as the modifications did not prevent target-specific cleavage of the target polynucleotide. In all the guide RNA sequences listed in the Table 1 of FIG. 4, the first 20 nucleotides at the 5' end are complementary to the target sequence in target DNA. The 40 modifications that were tested and found to be tolerated at specific positions include 2'-O-methylribonucleotide (=2'OMe), 2'-deoxyribonucleotide, racemic phosphorothioate internucleotide linkage(s) (=P(S)), 3'-phosphonoacetate (=PACE), 3'-thiophosphonoacetate (=thioPACE), Z nucleotides, and combinations of these.

It is contemplated that the chemical modifications disclosed and tested herein, particularly at the tested positions (as listed in the Table 1 of FIG. 4), will be tolerated at 50 equivalent positions in a variety of guide RNAs. In certain embodiments, the chemical modifications disclosed and tested herein are tolerated in any position in a guide RNA.

As disclosed herein, chemically modified nucleotides 55 were incorporated into guide RNAs in an effort to improve certain properties. Such properties include improved nucleic acid resistance of the guide RNA, reduced off-target effects of a gRNA:Cas protein complex (also known as improved specificity), improved efficacy of gRNA:Cas protein complex when cleaving, nicking or binding a target polynucleotide, improved transfection efficiency, and/or improved 60 organelle localization such as nuclear localization.

While the use of modified RNA is known (e.g., to block nucleolytic degradation in certain applications), it is widely known that one cannot simply incorporate modifications at any or all positions in an RNA sequence and expect it to function, particularly when the RNA sequence needs to complex with a protein or an enzyme to exert certain functions. Thus, it was not predictable whether these guide

US 10,337,001 B2

67

RNAs could tolerate chemical modifications at a variety of nucleotide positions while performing sufficient or improved function in a CRISPR-Cas system. In fact, it was unexpected that the guide RNA can tolerate specific modifications to the extent instantiated and tested, especially at several of the positions tested.

Example 2

To evaluate the ability of the chemically synthesized guide RNAs to target and cleave a DNA target sequence, an in vitro cleavage assay similar to that described in Example 1 was used. Target DNA constructs were for human DNA targets (sequences from the human clathrin light chain (CLTA1) gene, the human Interleukin 2 Receptor Gamma (IL2RG) gene, the human cytotoxic T-lymphocyte-associated protein 4 (CLTA4) gene, the human protocadherin alpha

4 (PCDHA4) gene, and the human engrailed homeobox 1 (EN1) gene), along with off-target DNA constructs differing from the target DNA by one or more nucleotides.

Table 3 sets forth the guide RNA constructs and their sequences, along with DNA constructs used for assessing the ability of those guide RNA constructs to target and cleave. In all the guide RNA sequences listed in the Table 3, the first 20 nucleotides at the 5' end are complementary to the target sequence in target DNA. ON target constructs comprise the 20 nt target sequence. OFF target constructs comprise most of the same 20 nucleotides as the target DNA, with 1, 2 or 3 nucleotide differences. Accordingly, the guide RNA is mostly, but not completely, complementary to the sequence of the OFF target constructs. The OFF target constructs are based on gene sequences known to occur in the human genome.

TABLE 3

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
2-piece dual-guide scaffold Unmodified dual-guide RNA (dgRNA)				
1	CLTA1 crRNA + tracrRNA	CLTA1 ON1-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC (SEQ ID NO: 25) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
2	CLTA1 crRNA + tracrRNA	CLTA1 ON1-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC (SEQ ID NO: 25) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
3	CLTA1 crRNA + tracrRNA	CLTA1 OFF1-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC (SEQ ID NO: 25) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
4	CLTA1 crRNA + tracrRNA	CLTA1 OFF1-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC + (SEQ ID NO: 25) GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
5	CLTA1 crRNA + tracrRNA	CLTA1 OFF2-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC (SEQ ID NO: 25) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
6	CLTA1 crRNA + tracrRNA	CLTA1 OFF2-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC (SEQ ID NO: 25) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
7	CLTA1 crRNA + tracrRNA	CLTA1 OFF3-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC (SEQ ID NO: 25) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
8	CLTA1 crRNA + tracrRNA	CLTA1 OFF3-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC (SEQ ID NO: 25) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
9	CLTA1 crRNA + tracrRNA	CLTA1 target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CAAAC (SEQ ID NO: 25) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 26)	56 + 86
10	IL2RG_crRNA + tracrRNA	IL2RGmg ON-target	UGGUAAAUGGCUUCAACAGUUUAAGAGCUAUGCUGUUUUGAAUGGUCC CCAAAAC (SEQ ID NO: 27) + GGAACCAUUCAAAACAGCAUAGCAAGGUUUAAAAGGCUGGUCCGUUAUCA ACUUGAAAAGUGGCCACCGAGUCGGUGGUUUUUU (SEQ ID NO: 28)	56 + 86

US 10,337,001 B2

69

70

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3') Fluorophore-coupled dgRNA	RNA length
Fluorophore-coupled dgRNA				
11	CLTA1 crRNA + tracrRNA aminoallyl- U57 + Cy5	CLTA1 ON1- target	AGUCUCUACUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUGAAUGGUCC CAAAC (SEQ ID NO: 29) + GGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUG (aminoallylU + Cy5) AAAAGUGGCACCGAGUCGGUGCUUUUU UU (SEQ ID NO: 30)	56 + 86
2'OMethyl-modified dgRNA				
12	IL2RG_crRNA_5',3'- 3x(2'OMe) + tracrRNA_5',3'- 3x(2'OMe)	IL2RGmg ON- target	U GGUAUGAUGGCUUCAACAGUUUAAGAGCUAUGCUGUUUGAAUGGU C CAAAAC (SEQ ID NO: 31) + G GAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 32)	56 + 86
2'OMethyl,3'Phosphorothioate-modified dgRNA				
13	IL2RG_crRNA_5',3'- 3x(2'OMe,3'P(S)) + tracrRNA_5',3'- 3x(2'OMe,3'P(S))	IL2RGmg ON- target	U sGsGsUAAGAUGGCUUCAACAGUUUAAGAGCUAUGCUGUUUGAAUGG UCCCA <u>sAsAs</u> (SEQ ID NO: 33) + G sGsAsACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUAU CAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU <u>sUsUs</u> (SEQ ID NO: 34)	56 + 86
2'OMethyl,3'PhosphorothioPACE-modified dgRNA				
14	IL2RG_crRNA_5',3'- 3x(2'OMe, 3'thioPACE) + tracrRNA_5',3'- 3x(2'OMe, 3'thioPACE)	IL2RGmg ON- target	U *sG*sG* s UAAGAUGGCUUCAACAGUUUAAGAGCUAUGCUGUUUGAAU GGUCCCA <u>sAsAs</u> (SEQ ID NO: 35) + G *sG*sA*sACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGU UAUCAACUUGAAAAGUGGCACCGAGUCGGUGCUUUU <u>sUsUs</u> (SEQ ID NO: 36)	56 + 86
15	IL2RG_crRNA_5',3'- 1x(2'OMe, 3'thioPACE) + tracrRNA_5',3'- 1x(2'OMe, 3'thioPACE)	IL2RGmg ON- target	U *sGGUAAGAUGGCUUCAACAGUUUAAGAGCUAUGCUGUUUGAAUGGU CCC <u>AAAAs</u> (SEQ ID NO: 37) + G *sGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAU CAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU <u>s</u> (SEQ ID NO: 38)	56 + 86
2-thioU-modified dgRNA				
16	CLTA1_2thioU + 3 crRNA + tracrRNA	CLTA1 ON1- target	AG (2sU) CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUGAAUGGU CCC <u>AAAAC</u> (SEQ ID NO: 39) + GGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 26)	56 + 86
17	CLTA1_2thioU + 3 crRNA + tracrRNA	CLTA1 ON1- target	AG (2sU) CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUGAAUGGU CCC <u>AAAAC</u> (SEQ ID NO: 39) + GGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 26)	56 + 86
18	CLTA1_2thioU + 3 crRNA + tracrRNA	CLTA1 OFF1- target	AG (2sU) CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUGAAUGGU CCC <u>AAAAC</u> (SEQ ID NO: 39) + GGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 26)	56 + 86
19	CLTA1_2thioU + 3 crRNA + tracrRNA	CLTA1 OFF1- target	AG (2sU) CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUGAAUGGU CCC <u>AAAAC</u> (SEQ ID NO: 39) + GGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 26)	56 + 86
20	CLTA1_2thioU + 3 crRNA + tracrRNA	CLTA1 OFF2- target	AG (2sU) CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUGAAUGGU CCC <u>AAAAC</u> (SEQ ID NO: 39) + GGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 26)	56 + 86
21	CLTA1_2thioU + 3 crRNA + tracrRNA	CLTA1 OFF2- target	AG (2sU) CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUGAAUGGU CCC <u>AAAAC</u> (SEQ ID NO: 39) + GGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 26)	56 + 86
22	CLTA1_2thioU + 3 crRNA + tracrRNA	CLTA1 OFF3- target	AG (2sU) CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGUUUGAAUGGU CCC <u>AAAAC</u> (SEQ ID NO: 39) + GGAACCAUAAAACAGCAUAGCAAGUUAAAUAAGGCUAUGUCCGUUAUCA ACUUGAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 26)	56 + 86

US 10,337,001 B2

71

72

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
23	CLTA1_2thioU + 3 crRNA + tracrRNA	CLTA1 OFF3-target	AG (2sU) CCCUCAUCUCCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 39) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
24	CLTA1_2thioU + 9 crRNA + tracrRNA	CLTA1 ON1-target	AGUCCUCA (2sU) CUCCCUCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 40) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
25	CLTA1_2thioU + 9 crRNA + tracrRNA	CLTA1 ON1-target	AGUCCUCA (2sU) CUCCCUCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 40) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
26	CLTA1_2thioU + 9 crRNA + tracrRNA	CLTA1 OFF1-target	AGUCCUCA (2sU) CUCCCUCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 40) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
27	CLTA1_2thioU + 9 crRNA + tracrRNA	CLTA1 OFF1-target	AGUCCUCA (2sU) CUCCCUCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 40) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
28	CLTA1_2thioU + 9 crRNA + tracrRNA	CLTA1 OFF2-target	AGUCCUCA (2sU) CUCCCUCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 40) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
29	CLTA1_2thioU + 9 crRNA + tracrRNA	CLTA1 OFF2-target	AGUCCUCA (2sU) CUCCCUCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 40) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
30	CLTA1_2thioU + 9 crRNA + tracrRNA	CLTA1 OFF3-target	AGUCCUCA (2sU) CUCCCUCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 40) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
31	CLTA1_2thioU + 9 crRNA + tracrRNA	CLTA1 OFF3-target	AGUCCUCA (2sU) CUCCCUCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 40) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
32	CLTA1_2thioU + 11 crRNA + tracrRNA	CLTA1 ON1-target	AGUCCUCAUC (2sU) CCCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 41) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
33	CLTA1_2thioU + 11 crRNA + tracrRNA	CLTA1 ON1-target	AGUCCUCAUC (2sU) CCCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 41) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
34	CLTA1_2thioU + 11 crRNA + tracrRNA	CLTA1 OFF1-target	AGUCCUCAUC (2sU) CCCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 41) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
35	CLTA1_2thioU + 11 crRNA + tracrRNA	CLTA1 OFF1-target	AGUCCUCAUC (2sU) CCCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 41) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
36	CLTA1_2thioU + 11 crRNA + tracrRNA	CLTA1 OFF2-target	AGUCCUCAUC (2sU) CCCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 41) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
37	CLTA1_2thioU + 11 crRNA + tracrRNA	CLTA1 OFF2-target	AGUCCUCAUC (2sU) CCCUCAAGCGUUUAAGAGCUAUGCUGUUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 41) + GGAACCAUCAAACAGCAUAGCAAGUUUUAAAAGGCCAUGUCCGUUAUCA ACUUGUAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86

US 10,337,001 B2

73

74

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
38	CLTA1_2thioU + 11 crRNA + tracrRNA	CLTA1 OFF3- target	AGUCUCUCAUC(2sU) CCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 41) + GGAACCAUCAAACAGCAUAGCAAGGUUUAAAAGGCCAUAGCUGGU ACUUGUAAAAGUGGCACCGACGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
39	CLTA1_2thioU + 11 crRNA + tracrRNA	CLTA1 OFF3- target	AGUCUCUCAUC(2sU) CCCUCAAGCGUUUAAGAGCUAUGCUGUUUUGAAUGGU CCCAAAAC (SEQ ID NO: 41) + GGAACCAUCAAACAGCAUAGCAAGGUUUAAAAGGCCAUAGCUGGU ACUUGUAAAAGUGGCACCGACGUCGGUGCUUUUUU (SEQ ID NO: 26)	56 + 86
Single-guide scaffold				
Unmodified single-guide RNA (sgRNA)				
40	CLTA1 sgRNA (Batch #1)	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCUAA6CGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
41	CLTA1 sgRNA (Batch #1)	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
42	CLTA1 sgRNA (Batch #2)	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
43	CLTA1 sgRNA (Batch #2)	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
44	CLTA1 sgRNA (Batch #3)	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
45	CLTA1 sgRNA (Batch #3)	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
46	CLTA1 SgRNA (Batch #3)	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
47	CLTA1 sgRNA (Batch #3)	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
48	CLTA1 sgRNA (Batch #3)	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
49	CLTA1 sgRNA (Batch #3)	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
50	CLTA1 sgRNA (Batch #3)	CLTA1mg OFF1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
51	CLTA1 sgRNA (Batch #3)	CLTA1mg OFF3- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
52	CLTA1 sgRNA (crude)	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 42)	113
53	CLTA1_Bos sgRNA	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 43)	100
54	CLTA1_Bos sgRNA	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 43)	100
55	CLTA1_Bos sgRNA	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCUAAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 43)	100

US 10,337,001 B2

75

76

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
56	CLTA1_Bos sgRNA	CLTA1mg OFF1- target	AGUCCUCAUCUCCCUAAGCGUUUUAGAGCUAGUAUAGCAAGGUAAAUA AGGUAGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGGGCUUU (SEQ ID NO: 43)	100
57	CLTA1_Bos sgRNA	CLTA1mg OFF3- target	AGUCCUCAUCUCCCUAAGCGUUUUAGAGCUAGUAUAGCAAGGUAAAUA AGGUAGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGGGCUUU (SEQ ID NO: 43)	100
58	CLTA4_SgRNA	CLTA4 ON-target	GCAGAUGUAUGGUUUCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 44)	113
59	CLTA4_sgRNA	CLTA4 ON-target	GCAGAUGUAUGGUUUCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 44)	113
60	CLTA4_sgRNA	CLTA4 ON-target	GCAGAUGUAUGGUUUCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 44)	113
61	CLTA4_sgRNA	CLTA4mg ON-target	GCAGAUGUAUGGUUUCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 44)	113
62	CLTA4_sgRNA	CLTA4mg ON-target	GCAGAUGUAUGGUUUCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 44)	113
63	CLTA4_sgRNA	CLTA4mg OFF5- target	GCAGAUGUAUGGUUUCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 44)	113
64	CLTA1_Truncated_18mer	CLTA1mg ON1- target	UCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAG UUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGG UGCUUUUUUU (SEQ ID NO: 45)	111
65	CLTA1_Truncated_18mer	CLTA1mg ON1- target	UCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAG UUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGG UGCUUUUUUU (SEQ ID NO: 45)	111
66	CLTA1_Truncated_18mer	CLTA1mg OFF1- target	UCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAG UUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGG UGCUUUUUUU (SEQ ID NO: 45)	111
67	CLTA1_Truncated_18mer	CLTA1mg OFF3- target	UCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAG UUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGG UGCUUUUUUU (SEQ ID NO: 45)	111
68	CLTA1_Truncated_17mer	CLTA1mg ON1- target	CCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAGU UUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGG GCUUUUUUU (SEQ ID NO: 46)	110
69	CLTA1_Truncated_17mer	CLTA1mg ON1- target	CCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAGU UUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGG GCUUUUUUU (SEQ ID NO: 46)	110
70	CLTA1_Truncated_17mer	CLTA1mg OFF1- target	CCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAGU UUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGG GCUUUUUUU (SEQ ID NO: 46)	110
71	CLTA1_Truncated_17mer	CLTA1mg OFF3- target	CCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAGU UUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGUCGG GCUUUUUUU (SEQ ID NO: 46)	110
72	CLTA1_1xExtraG	CLTA1mg ON1- target	GAGUCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 47)	114
73	CLTA1_1xExtraG	CLTA1mg ON1- target	GAGUCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 47)	114
74	CLTA1_1xExtraG	CLTA1mg OFF1- target	GAGUCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGCUAUGUCCGUUAUACAUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 47)	114

US 10,337,001 B2

77

78

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
75	CLTA1_1xExtraG	CLTA1mg OFF3-target	GAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG CGGUGCUUUUUU (SEQ ID NO: 47)	114
76	CLTA1_2xExtraG	CLTA1mg ON1-target	GGAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUUUU (SEQ ID NO: 48)	115
77	CLTA1_2xExtraG	CLTA1mg ON1-target	GGAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUUUU (SEQ ID NO: 48)	115
78	CLTA1_2xExtraG	CLTA1mg OFF1-target	GGAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUUUU (SEQ ID NO: 48)	115
79	CLTA1_2xExtraG	CLTA1mg OFF3-target	GGAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUUUU (SEQ ID NO: 48)	115
80	CLTA1_63U, 64U	CLTA1mg ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAUUCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUU (SEQ ID NO: 49)	113
81	CLTA1_63A, 64A	CLTA1mg ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAACUAGUCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUU (SEQ ID NO: 50)	113
82	CLTA1_63A, 64A, 70U, 71U	CLTA1mg ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAACUAGUCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUU (SEQ ID NO: 51)	113
83	CLTA1_cis-block(1- 5)_polyU_sgRNA	CLTA1mg ON1-target	GGACUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGAAAAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU (SEQ ID NO: 52)	111
84	CLTA1_cis-block(1- 5)_polyU_sgRNA	CLTA1mg ON1-target	GGACUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGAAAAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU (SEQ ID NO: 52)	111
85	CLTA1_cis-block(1- 5)_polyU_sgRNA	CLTA1mg OFF1-target	GGACUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGAAAAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU (SEQ ID NO: 52)	111
86	CLTA1_cis-block(1- 5)_polyU_sgRNA	CLTA1mg OFF3-target	GGACUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGAAAAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU (SEQ ID NO: 52)	111
87	CLTA1_cis-block(1- 10)_polyU_sgRNA	CLTA1mg ON1-target	GAUGAGGACUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGA AAUAGCAAGUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG CCGAGUCGGUGCUUUU (SEQ ID NO: 53)	116
88	CLTA1_cis-block(1- 10)_polyU_sgRNA	CLTA1mg ON1-target	GAUGAGGACUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGA AAUAGCAAGUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG CCGAGUCGGUGCUUUU (SEQ ID NO: 53)	116
89	CLTA1_cis-block(1- 10)_polyU_sgRNA	CLTA1mg OFF1-target	GAUGAGGACUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGA AAUAGCAAGUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG CCGAGUCGGUGCUUUU (SEQ ID NO: 53)	116
90	CLTA1_cis-block(1- 10)_polyU_sgRNA	CLTA1mg OFF3-target	GAUGAGGACUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGA AAUAGCAAGUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG CCGAGUCGGUGCUUUU (SEQ ID NO: 53)	116
91	CLTA1_cis-block(16- 20)_polyU_sgRNA	CLTA1mg ON1-target	GCUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGAAAAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU (SEQ ID NO: 54)	111
92	CLTA1_cis-block(16- 20)_polyU_sgRNA	CLTA1mg ON1-target	GCUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGAAAAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU (SEQ ID NO: 54)	111
93	CLTA1_cis-block(16- 20)_polyU_sgRNA	CLTA1mg OFF1-target	GCUUUUUUUAGGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAGAAAAG CAAGUUUUAAAAGGCUAGUCGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU (SEQ ID NO: 54)	111

US 10,337,001 B2

79

80

TABLE 3-continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
94	CLTA1_cis-block(16-20)_polyU_sgRNA	CLTA1mg OFF3-target	GC芋GUUUUUAGUCUCAUCCCAAGCGUUUAAGAGCUAGAAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAAGUGGCACCGAGUCGGUCUUUU (SEQ ID NO: 54)	111
DMT-modified sgRNA				
95	CLTA1_DMT-ON SgRNA	CLTA1 ON1-target	(dmt) AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 55)	113
96	CLTA1_DMT-ON/ OFF sgRNA	CLTA1 ON1-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 56)	113
Fluorophore-modified sgRNA				
97	CLTA1_IntF1_sgLoop	CLTA1 ON1-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 57)	113
98	CLTA1_IntF1_sgLoop	CLTA1mg ON1-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 57)	113
99	CLTA1_IntF1_sgLoop	CLTA1mg ON1-target	AGUCCUCAUCUCCCUCAA6CGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 57)	113
100	CLTA1_IntF1_sgLoop	CLTA1mg OFF3-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 57)	113
101	CLTA1_IntF1_sgLoop	CLTA1mg OFF3-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 57)	113
102	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe)	CLTA1mg ON1-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 58)	113
103	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe)	CLTA1mg ON1-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 58)	113
104	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe)	CLTA1mg OFF3-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 58)	113
105	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe)	CLTA1mg OFF3-target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO: 58)	113
106	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe, 3'P(S))	CLTA1mg ON1-target	AsGsUsCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUsUsUs (SEQ ID NO: 59)	113
107	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe, 3'P(S))	CLTA1mg ON1-target	AsGsUsCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUsUsUs (SEQ ID NO: 59)	113
108	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe, 3'P(S))	CLTA1mg OFF3-target	AsGsUsCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUsUsUs (SEQ ID NO: 59)	113
109	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe, 3'P(S))	CLTA1mg OFF3-target	AsGsUsCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGAGUCGGUGCUUUUsUsUs (SEQ ID NO: 59)	113
110	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA1mg ON1-target	AsGsUs*SCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGACCGAGUCGGUGCUUUUsUsUs (SEQ ID NO: 60)	113
111	CLTA1_IntF1_sgLoop_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA1mg ON1-target	A*sG*sU*SCCUCAUCUCCCUAAGCGUUUAAGAGCUAUGCUGG(F1) AACAGCAUAGCAAGUUAAAUAAGGCUAUGCCGUUAACUUGAAAAGUGGCACCGACCGAGUCGGUGCUUUUsUsUs (SEQ ID NO: 60)	113

US 10,337,001 B2

81

82

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
112	CLTA1_1ntFl_sgLoop_5', 3'- 3x(2'OMe, 3'thioPACE)	CLTA1mg OFF1- target	A*sG*sU*sCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGG (F1) AACAA GCAUAGCAAGUUAAAAGCUAGUCGUUAUCAACUUGAAAAAGUGGCAC CGAGUCGGUCUUU <u>sU*sU*sU</u> (SEQ ID NO: 60)	113
113	CLTA1_1ntFl_sgLoop_5', 3'- 3x(2'OMe, 3'thioPACE)	CLTA1mg OFF3- target	A*sG*sU*sCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGG (F1) AACAA GCAUAGCAAGUUAAAAGCUAGUCGUUAUCAACUUGAAAAAGUGGCAC CGAGUCGGUCUUU <u>sU*sU*sU</u> (SEQ ID NO: 60)	113
114	CLTA4_3xF1- Int_3x(2'OMe, 3'thioPACE)	CLTA4mg ON- target	U,*s (F1,) GCAGAUGUAGGUUUCACAGUUUAAGAGCUAUGAACAGU UuAAAUAAGGUAGUCGUUA (F1) CAACUUGAAAAGUGGCACCGAGUCGG UGCU (F1) U*sU (SEQ ID NO: 61)	102
115	CLTA4_3xF1- Int_3x(2'OMe, 3'thioPACE)	CLTA4mg OFF5- target	U,*s (F1,) GCAGAUGUAGGUUUCACAGUUUAAGAGCUAUGAACAGU UuAAAUAAGGUAGUCGUUA (F1) CAACUUGAAAAGUGGCACCGAGUCGG UGCU (F1) U*sU (SEQ ID NO: 61)	102
116	CLTA4_3xF1- Loops_3x(2'OMe, 3'thioPACE)	CLTA4mg ON- target	G*sCAGAUGUAGGUUUCACAGUUUAAGAGCUAUG (F1) AAUAGCAAGUU AAAUAAGGUAGUCGUUAUCAACUUG (F1) AAAAGUGGCACCGAG (F1) C GGUGCUUU <u>sU</u> (SEQ ID NO: 62)	100
117	CLTA4_3xF1- Loops_3x(2'OMe, 3'thioPACE)	CLTA4mg OFF5- target	G*sCAGAUGUAGGUUUCACAGUUUAAGAGCUAUG (F1) AAUAGCAAGUU AAAUAAGGUAGUCGUUAUCAACUUG (F1) AAAAGUGGCACCGAG (F1) C GGUGCUUU <u>sU</u> (SEQ ID NO: 62)	100
3' Phosphorothioate-modified sgRNA				
118	CLTA1_5'-2xP(S) sgRNA	CLTA1 ON1- target	AsGsUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 63)	113
119	CLTA1_5'-3xP(S) sgRNA	CLTA1 ON1- target	AsGsUsSCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGCUUUUUUU (SEQ ID NO: 64)	113
120	CLTA1_5'-4xP(S) sgRNA	CLTA1 ON1- target	AsGsUsCsCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGA GUCGGUGCUUUUUUU (SEQ ID NO: 65)	113
121	CLTA1_3'-4xP(S) SgRNA	CLTA1 ON1- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUsUsUsU (SEQ ID NO: 66)	113
2'OMethyl-modified sgRNA				
122	CLTA1_2'OMe + 20 sgRNA	CLTA1 ON1- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 67)	113
123	CLTA1_2'OMe + 19 sgRNA	CLTA1 ON1- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 68)	113
124	CLTA1_2'OMe + 19 sgRNA	CLTA1mg ON1- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 68)	113
125	CLTA1_2'OMe + 19 sgRNA	CLTA1mg ON1- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 68)	113
126	CLTA1_2'OMe + 19 sgRNA	CLTA1mg OFF1- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 68)	113
127	CLTA1_2'OMe + 19 sgRNA	CLTA1mg OFF3- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 68)	113
128	CLTA1_2'OMe + 18 sgRNA	CLTA1 ON1- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 69)	113
129	CLTA1_2'OMe + 18 sgRNA	CLTA1mg ON1- target	AGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 69)	113

US 10,337,001 B2

83

84

TABLE 3-continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
130	CLTA1_2'OMe + 18 sgRNA	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 69)	113
131	CLTA1_2'OMe + 18 sgRNA	CLTA1mg OFF1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 69)	113
132	CLTA1_2'OMe + 18 sgRNA	CLTA1mg OFF3- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 69)	113
133	CLTA1_2'OMe + 17 sgRNA	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 70)	113
134	CLTA1_2'OMe + 17 sgRNA	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 70)	113
135	CLTA1_2'OMe + 17 sgRNA	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 70)	113
136	CLTA1_2'OMe + 17 sgRNA	CLTA1mg OFF1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 70)	113
137	CLTA1_2'OMe + 17 sgRNA	CLTA1mg OFF3- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 70)	113
138	CLTA1_2'OMe + 17, 18 sgRNA	CLTA1 ON1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 71)	113
139	CLTA1_2'OMe + 17, 18 sgRNA	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 71)	113
140	CLTA1_2'OMe + 17, 18 sgRNA	CLTA1mg ON1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 71)	113
141	CLTA1_2'OMe + 17, 18 sgRNA	CLTA1mg OFF1- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 71)	113
142	CLTA1_2'OMe + 17, 18 sgRNA	CLTA1mg OFF3- target	AGUCUCUCAUCUCCCUCA <u>A</u> GCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 71)	113
143	CLTA1_5', 3'- 3x(2'OMe) sgRNA	CLTA1 ON1- target	<u>A</u> GUCUCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 72)	113
144	CLTA4_5', 3'- 3x(2'OMe) sgRNA	CLTA4mg ON- target	<u>G</u> CAGAUGAUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 73)	113
145	CLTA4_5', 3'- 3x(2'OMe) sgRNA	CLTA4mg OFF5- target	<u>G</u> CAGAUGAUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 73)	113
146	CLTA1_5'- 20x(2'OMe) sgRNA	CLTA1 ON1- target	<u>A</u> GUCUCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 74)	113
147	CLTA1_5'- 20x(2'OMe) sgRNA	CLTA1mg ON1- target	<u>A</u> GUCUCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 74)	113
148	CLTA1_5'- 20x(2'OMe) sgRNA	CLTA1mg ON1- target	<u>A</u> GUCUCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 74)	113

US 10,337,001 B2

85

86

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
149	CLTA1_5'- 20x(2'OMe) sgRNA	CLTA1mg OFF1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 74)	113
150	CLTA1_5'- 20x(2'OMe) sgRNA	CLTA1mg OFF3-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 74)	113
151	CLTA1_5'- 26x(2'OMe) sgRNA	CLTA1 ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AAGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC CGGUGCUUUUUUU (SEQ ID NO: 75)	113
152	CLTA1_5'- 37x(2'OMe) SgRNA	CLTA1 ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AAGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC CGGUGCUUUUUUU (SEQ ID NO: 76)	113
153	CLTA1_41x(2'OMeC/U)_QB3	CLTA1mg ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 77)	113
154	CLTA1_47x(2'OMeC/U)_QB3	CLTA1 ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 78)	113
155	CLTA1_47x(2'OMeC/U)_QB3	CLTA1mg ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 79)	113
156	CLTA1_47x(2'OMeC/U)_QB3	CLTA1mg ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 79)	113
157	CLTA1_47x(2'OMeC/U)_QB3	CLTA1mg OFF1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 79)	113
158	CLTA1_47x(2'OMeC/U)_QB3	CLTA1mg OFF3-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 79)	113
159	CLTA1_47x(2'OMeG/A)_QB3	CLTA1 ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 80)	113
160	CLTA1_47x(2'OMeG/A)_QB3	CLTA1mg ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 80)	113
161	CLTA1_47x(2'OMeG/A)_QB3	CLTA1mg ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 80)	113
162	CLTA1_47x(2'OMeG/A)_QB3	CLTA1mg OFF1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 80)	113
163	CLTA1_47x(2'OMeG/A)_QB3	CLTA1mg OFF3-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGUUAAAUAAGGCUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUC GGUGCUUUUUUU (SEQ ID NO: 80)	113
164	CLTA1_43x(2'OMeG/A)_Bos	CLTA1 ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGGUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 81)	100
165	CLTA1_43x(2'OMeG/A)_Bos	CLTA1mg ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGGUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 81)	100
166	CLTA1_43x(2'OMeG/A)_Bos	CLTA1mg ON1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGGUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 81)	100
167	CLTA1_43x(2'OMeG/A)_Bos	CLTA1mg OFF1-target	<u>AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA</u> AGGUAGCCGUUAUCACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 81)	100

US 10,337,001 B2

87

88

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
168	CLTA1_43x(2'OMeG/A)_Bos	CLTA1mg OFF3 - target	<u>AGUCCUCAUCCCCUCAAGC</u> <u>GUUUUAGAGCUAUGUAUAGCAAGUAAAUA</u> <u>AGGCUAUGC</u> <u>GGUUAUCAACUUGAAAAGUGGCACCGAGUCGG</u> <u>GUCGUUUU</u> (SEQ ID NO: 81)	100
169	CLTA4_sgRNA_5',3'-3x(2'OMe)	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UGGAAACAGCAUAGC</u> <u>AAGUUUAAAUAAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGU</u> <u>CGGUGCUUUUUUU</u> (SEQ ID NO: 82)	113
170	CLTA4_sgRNA_5',3'-3x(2'OMe)	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UGGAAACAGCAUAGC</u> <u>AAGUUUAAAUAAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGU</u> <u>CGGUGCUUUUUUU</u> (SEQ ID NO: 82)	113
171	CLTA4_47x(2'OMeC/U)_QB3	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UGGUAACAGCAUAGC</u> <u>AAGUUUAAAUAAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGU</u> <u>CGGUGCUUUUUUU</u> (SEQ ID NO: 83)	113
172	CLTA4_47x(2'OMeC/U)_QB3	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UGGUAACAGCAUAGC</u> <u>AAGUUUAAAUAAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGU</u> <u>CGGUGCUUUUUUU</u> (SEQ ID NO: 83)	113
173	CLTA4_47x(2'OMeC/U)_QB3	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UGGUAACAGCAUAGC</u> <u>AAGUUUAAAUAAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGU</u> <u>CGGUGCUUUUUUU</u> (SEQ ID NO: 83)	113
174	CLTA4_49x(2'OMeG/A)_Bos	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UAUGCAAGUAAAUA</u> <u>AAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGUCGG</u> <u>UUCUUU</u> (SEQ ID NO: 84)	100
175	CLTA4_49x(2'OMeG/A)_Bos	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UAUGCAAGUAAAUA</u> <u>AAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGUCGG</u> <u>UUCUUU</u> (SEQ ID NO: 84)	100
176	CLTA4_49x(2'OMeG/A)_Bos	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UAUGCAAGUAAAUA</u> <u>AAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGUCGG</u> <u>UUCUUU</u> (SEQ ID NO: 84)	100
177	CLTA4_39x(2'OMeC/U)_BoS	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UAUGCAAGUAAAUA</u> <u>AAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGUCGG</u> <u>UUCUUU</u> (SEQ ID NO: 85)	100
178	CLTA4_39x(2'OMeC/U)_BoS	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UAUGCAAGUAAAUA</u> <u>AAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGUCGG</u> <u>UUCUUU</u> (SEQ ID NO: 85)	100
179	CLTA4_39x(2'OMeC/U)_BoS	CLTA4 ON-target	<u>GCAGAUGUAGUGUUUCCACAGUUUAGAGCUAUGC</u> <u>UAUGCAAGUAAAUA</u> <u>AAGGC</u> <u>UAGGUAGUCCG</u> <u>UUAUCAACUUGAAAAGUGGCACCGAGUCGG</u> <u>UUCUUU</u> (SEQ ID NO: 85)	100
2'Deoxy-modified sgRNA				
180	CLTA1_5'-20x(2'deoxy)	CLTA1 ON1-target	<u>AGTCCTCATCTCCCTCAAGC</u> <u>UUUAAGAGCUAUGC</u> <u>UGGUAAACAGCAUAG</u> <u>CAAGUUAAAAGCUAGUCC</u> <u>GUUAUCAACUUGAAAAGUGGCACCGAG</u> <u>UCGGUGCUUUUUUU</u> (SEQ ID NO: 86)	113
181	CLTA1_5'-26x(2'deoxy)	CLTA1 ON1-target	<u>AGTCCTCATCTCCCTCAAGC</u> <u>GTTTAA</u> <u>GAGCUAUGC</u> <u>UGGUAAACAGCAUAG</u> <u>CAAGUUAAAAGCUAGUCC</u> <u>GUUAUCAACUUGAAAAGUGGCACCGAG</u> <u>CGGUGCUUUUUUU</u> (SEQ ID NO: 87)	113
182	CLTA1_5'-37x(2'deoxy)	CLTA1 ON1-target	<u>AGTCCTCATCTCCCTCAAGC</u> <u>GTTTAA</u> <u>GAGCTATGCTG</u> <u>UAACAGCAUA</u> <u>GCAAGUUAAAAGCUAGUCC</u> <u>GUUAUCAACUUGAAAAGUGGCACCGAG</u> <u>UCGGUGCUUUUUUU</u> (SEQ ID NO: 88)	113
2'Deoxy, 3'PACE-modified sgRNA				
183	CLTA4_2'deoxy3'PACE + 15	CLTA4mg ON-target	<u>GCAGAUGUAGUGU</u> <u>U*</u> <u>CCACAGUUUAGAGCUAUGC</u> <u>UGGUAAACAGCAUAG</u> <u>CAAGUUAAAAGCUAGUCC</u> <u>GUUAUCAACUUGAAAAGUGGCACCGAG</u> <u>UCGGUGCUUUUUUU</u> (SEQ ID NO: 89)	113
184	CLTA4_2'deoxy3'PACE + 15	CLTA4mg OFF5-target	<u>GCAGAUGUAGUGU</u> <u>U*</u> <u>CCACAGUUUAGAGCUAUGC</u> <u>UGGUAAACAGCAUAG</u> <u>CAAGUUAAAAGCUAGUCC</u> <u>GUUAUCAACUUGAAAAGUGGCACCGAG</u> <u>UCGGUGCUUUUUUU</u> (SEQ ID NO: 89)	113

US 10,337,001 B2

89

90

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3') 2'OMethyl, 3'PACE-modified sgRNA	RNA length
185	5'-1x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	A*GUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGC AAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 90)	113
186	5'-1x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	A*GUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGC AAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 90)	113
187	5'-2x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1 ON1-target	A*G*UCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGC CAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU UCGGUGCUUUUUU (SEQ ID NO: 91)	113
188	5'-2x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1 ON1-target	A*G*UCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGC CAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU UCGGUGCUUUUUU (SEQ ID NO: 91)	113
189	5'-2x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	A*G*UCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGC CAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU UCGGUGCUUUUUU (SEQ ID NO: 91)	113
190	5'-2x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	A*G*UCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGC CAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU UCGGUGCUUUUUU (SEQ ID NO: 91)	113
191	5'-3x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	G*G*A*GUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCA UAGCAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCG GAGUCGGUGCUUUUUU (SEQ ID NO: 92)	115
192	5'-3x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	G*G*A*GUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCA UAGCAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCG GAGUCGGUGCUUUUUU (SEQ ID NO: 92)	115
193	5'-4x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1 ON1-target	A*G*U*C*CUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCA AGCAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 93)	113
194	5'-4x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	A*G*U*C*CUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCA AGCAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 93)	113
195	5'-4x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	A*G*U*C*CUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCA AGCAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 93)	113
196	5'-5x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	G*G*A*G*U*CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAG CAUAGCAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCG CCGAGUCGGUGCUUUUUU (SEQ ID NO: 94)	115
197	5'-5x(2'OMe, 3'PACE)_CLTA1 sgRNA	CLTA1mg ON1-target	G*G*A*G*U*CCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAG CAUAGCAAGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCG CCGAGUCGGUGCUUUUUU (SEQ ID NO: 94)	115
198	CLTA1_3'- 4x(2'OMe, 3'PACE) sgRNA	CLTA1 ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGCA AGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUU*U*U*U*U (SEQ ID NO: 95)	113
199	CLTA1_3'- 4x(2'OMe, 3'PACE) sgRNA	CLTA1 ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGCA AGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUU*U*U*U*U (SEQ ID NO: 95)	113
200	CLTA1_3'- 4x(2'OMe, 3'PACE) sgRNA	CLTA1mg ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGCA AGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUU*U*U*U*U (SEQ ID NO: 95)	113
201	CLTA1_3'- 4x(2'OMe, 3'PACE) sgRNA	CLTA1mg ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGCA AGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUU*U*U*U*U (SEQ ID NO: 95)	113
202	CLTA1_3'- 5x(2'OMe, 3'PACE) sgRNA	CLTA1 ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGCA AGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUU*U*U*U*U (SEQ ID NO: 96)	113
203	CLTA1_3'- 5x(2'OMe, 3'PACE) sgRNA	CLTA1 ON1-target	AGUCCUCAUCUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUAAACAGCAUAGCA AGUUUUAAAAGGCCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUU*U*U*U*U (SEQ ID NO: 96)	113

US 10,337,001 B2

91

92

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
204	CLTA1_3'- 5x(2'OMe, 3'PACE) sgRNA	CLTA1mg ON1- target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGCACCGAGUC GGUGCUU*U*U*U*U*U (SEQ ID NO: 96)	113
205	CLTA1_3'- 5x(2'OMe, 3'PACE) sgRNA	CLTA1mg ON1- target	AGUCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGCACCGAGUC GGUGCUU*U*U*U*U*U (SEQ ID NO: 96)	113
206	5'-3x(2'OMe, 3'PACE)_plus1 overhg_CLTA1	CLTA1 ON1- target	C,*A*G*UCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 97)	114
207	5'-3x(2'OMe, 3'PACE)_plus1 overhg_CLTA1	CLTA1mg ON1- target	C,*A*G*UCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 97)	114
208	5'-3x(2'OMe, 3'PACE)_plus1 overhg_CLTA1	CLTA1mg ON1- target	C,*A*G*UCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 97)	114
209	5'-3x(2'OMe, 3'PACE)_plus1 NC_overhg_CLTA1	CLTA1 ON1- target	G,*A*G*UCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 98)	114
210	5'-3x(2'OMe, 3'PACE)_plus1 NC_overhg_CLTA1	CLTA1mg ON1- target	G,*A*G*UCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 98)	114
211	5'-3x(2'OMe, 3'PACE)_plus1 NC_overhg_CLTA1	CLTA1mg ON1- target	G,*A*G*UCCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 98)	114
212	5'-5x(2'OMe, 3'PACE)_plus2 overhg_CLTA1	CLTA1 ON1- target	U,*C,*A*G*U*CCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUAC GCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGC ACCGAGUCGGUGCUUUUUU (SEQ ID NO: 99)	115
213	5'-5x(2'OMe, 3'PACE)_plus2 overhg_CLTA1	CLTA1mg ON1- target	U,*C,*A*G*U*CCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUAC GCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGC ACCGAGUCGGUGCUUUUUU (SEQ ID NO: 99)	115
214	5'-5x(2'OMe, 3'PACE)_plus2 overhg_CLTA1	CLTA1mg ON1- target	U,*C,*A*G*U*CCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUAC GCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGC ACCGAGUCGGUGCUUUUUU (SEQ ID NO: 99)	115
215	5'-5x(2'OMe, 3'PACE)_plus2 NC_overhg_CLTA1	CLTA1 ON1- target	A,*G,*A*G*U*CCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACA GCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGC ACCGAGUCGGUGCUUUUUU (SEQ ID NO: 100)	115
216	5'-5x(2'OMe, 3'PACE)_plus2 NC_overhg_CLTA1	CLTA1mg ON1- target	A,*G,*A*G*U*CCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACA GCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGC ACCGAGUCGGUGCUUUUUU (SEQ ID NO: 100)	115
217	5'-5x(2'OMe, 3'PACE)_plus2 NC_overhg_CLTA1	CLTA1mg ON1- target	A,*G,*A*G*U*CCUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGUACA GCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUGGC ACCGAGUCGGUGCUUUUUU (SEQ ID NO: 100)	115
218	5'-7x(2'OMe, 3'PACE)_plus3 overhg_CLTA1_3'- 4x(2'OMe, 3'PACE)	CLTA1 ON1- target	C,*U,*C,*A*G*U*C*CUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGU ACAGCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUG GGCACCGAGUCGGUGCUUU*U*U*U*U (SEQ ID NO: 101)	116
219	5'-7x(2'OMe, 3'PACE)_plus3 overhg_CLTA1_3'- 4x(2'OMe, 3'PACE)	CLTA1mg ON1- target	C,*U,*C,*A*G*U*C*CUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGU ACAGCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUG GGCACCGAGUCGGUGCUUU*U*U*U*U (SEQ ID NO: 101)	116
220	5'-7x(2'OMe, 3'PACE)_plus3 overhg_CLTA1_3'- 4x(2'OMe, 3'PACE)	CLTA1mg ON1- target	C,*U,*C,*A*G*U*C*CUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGU ACAGCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUG GGCACCGAGUCGGUGCUUU*U*U*U*U (SEQ ID NO: 101)	116
221	5'-7x(2'OMe, 3'PACE)_plus3 NC_overhg_CLTA1_3'- 4x(2'OMe, 3'PACE)	CLTA1 ON1- target	G,*A,*G,*A*G*U*C*CUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGU ACAGCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUG GGCACCGAGUCGGUGCUUU*U*U*U*U (SEQ ID NO: 101)	116
222	5'-7x(2'OMe, 3'PACE)_plus3 NC_overhg_CLTA1_3'- 4x(2'OMe, 3'PACE)	CLTA1mg ON1- target	G,*A,*G,*A*G*U*C*CUCAUCUCCCUCAAGCGUUUAAGAGCUAUGCUGGU ACAGCAUAGCAAGUUAAAUAAGGCUGUCCGUUAUCACUUGAAAAGUG GGCACCGAGUCGGUGCUUU*U*U*U*U (SEQ ID NO: 101)	116

US 10,337,001 B2

93

94

TABLE 3-continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
223	5'-7x(2'OMe, 3'PACE)_plus3 NC_overhg_CLTA1_3'-4x(2'OMe, 3'PACE)	CLTA1mg ON1-target	G ₁ A ₂ *G ₃ *A ₄ G ₅ *U ₆ C ₇ *CUCAUCUCCCUCAGGUUUAGAGCUAUGCUGGUACAGCAUAGC ACAGCAUAGCAAGGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGGU GGCACCGAGUCGGUCUUUU <u>U</u> * <u>U</u> <u>U</u> * <u>U</u> (SEQ ID NO: 101)	116
224	CLTA1_2'OMe, 3'PACE + 20 sgRNA	CLTA1 ON1-target	AGUCCUCAUCCCCUCAAGC*GUUUAAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 102)	113
225	CLTA1_2'OMe, 3'PACE + 20 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAGC*GUUUAAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 102)	113
226	CLTA1_2'OMe, 3'PACE + 20 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAGC*GUUUAAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 102)	113
227	CLTA1_2'OMePACE + 19 sgRNA	CLTA1 ON1-target	AGUCCUCAUCCCCUCAAG*CGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 103)	113
228	CLTA1_2'OMePACE + 19 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAG*CGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 103)	113
229	CLTA1_2'OMePACE + 19 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAG*CGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 103)	113
230	CLTA1_2'OMePACE + 19 sgRNA	CLTA1mg OFF1-target	AGUCCUCAUCCCCUCAAG*CGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 103)	113
231	CLTA1_2'OMePACE + 19 sgRNA	CLTA1mg OFF3-target	AGUCCUCAUCCCCUCAAG*CGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 103)	113
232	CLTA1_2'OMePACE + 18 sgRNA	CLTA1 ON1-target	AGUCCUCAUCCCCUCAAG*GCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 104)	113
233	CLTA1_2'OMePACE + 18 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAG*GCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 104)	113
234	CLTA1_2'OMePACE + 18 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAG*GCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 104)	113
235	CLTA1_2'OMePACE + 17 sgRNA	CLTA1 ON1-target	AGUCCUCAUCCCCUCAAG*AGCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 105)	113
236	CLTA1_2'OMePACE + 17 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAG*AGCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 105)	113
237	CLTA1_2'OMePACE + 17 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAG*AGCGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAGU CGGUGCUUUUUUU (SEQ ID NO: 105)	113
238	CLTA1_2'OMePACE + 17, 18 sgRNA	CLTA1 ON1-target	AGUCCUCAUCCCCUCAAG*GCGUUUAAGAGCUAUGCUGGUACAGCAUAGC CAAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUUUUU (SEQ ID NO: 106)	113
239	CLTA1_2'OMePACE + 17, 18 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAG*GCGUUUAAGAGCUAUGCUGGUACAGCAUAGC CAAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUUUUU (SEQ ID NO: 106)	113
240	CLTA1_2'OMePACE + 17, 18 sgRNA	CLTA1mg ON1-target	AGUCCUCAUCCCCUCAAG*GCGUUUAAGAGCUAUGCUGGUACAGCAUAGC CAAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUUUUUU (SEQ ID NO: 106)	113
2'OMethyl, 3'Phosphorothioate-modified sgRNA				
241	CLTA1_5', 3'-3x(2'OMe, 3'P(S))	CLTA1 ON1-target	AsGsUs ₁ CCUCAUCUCCCCUCAAGGUUUAGAGCUAUGCUGGUACAGCAUAGC CAAGUUUAAAUAAGGUAGUCGGUUAUCACUUGAAAAAGUGGCACCGAG UCGGUGCUUU <u>U</u> <u>s</u> <u>U</u> <u>s</u> <u>U</u> (SEQ ID NO: 107)	113

US 10,337,001 B2

95

96

TABLE 3-continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
242	CLTA1_5',3'- 3x(2'OMe, 3'P(S))	CLTA1mg ON1- target	<u>AsGsUs</u> CCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGAG UCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
243	CLTA1_5',3'- 3x(2'OMe, 3'P(S))	CLTA1mg ON1- target	<u>AsGsUs</u> CCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGAG UCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
244	CLTA1_5',3'- 3x(2'OMe, 3'P(S))	CLTA1mg OFF1- target	<u>AsGsUs</u> CCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGAG UCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
245	CLTA1_5',3'- 3x(2'OMe, 3'P(S))	CLTA1mg OFF3- target	<u>AsGsUs</u> CCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGAG UCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
246	CLTA4_5',3'- 3x(2'OMe, 3'P(S))	CLTA4 target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
247	CLTA4_5',3'- 3x(2'OMe, 3'P(S))	CLTA4 target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
248	CLTA4_5',3'- 3x(2'OMe, 3'P(S))	CLTA4mg ON- target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
249	CLTA4_5',3'- 3x(2'OMe, 3'P(S))	CLTA4mg ON- target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
250	CLTA4_5',3'- 3X(2'OMe, 3'P(S))	CLTA4mg ON- target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
251	CLTA4_5',3'- 3x(2'OMe, 3'P(S))	CLTA4mg OFF5- target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsU</u> (SEQ ID NO: 107)	113
252	CLTA4_5'- 3x(2'OMe, 3'P(S)),3'- 5x(2'OMe, 3'P(S))	CLTA4mg ON- target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsUsU</u> (SEQ ID NO: 108)	113
253	CLTA4_5'- 3x(2'OMe, 3'P(S)),3'- 5x(2'OMe, 3'P(S))	CLTA4mg ON- target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsUsU</u> (SEQ ID NO: 108)	113
254	CLTA4_5'- 3x(2'OMe, 3'P(S)),3'- 5x(2'OMe, 3'P(S))	CLTA4mg ON- target	<u>GsCsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAUA GCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCGA GUCGGUGCUUU <u>sUsUsUsU</u> (SEQ ID NO: 108)	113
255	CLTA4_5',3'- 5x(2'OMe, 3'P(S))	CLTA4mg ON- target	<u>GsCsAsGsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAU AGCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUU <u>sUsUsUsU</u> (SEQ ID NO: 109)	113
256	CLTA4_5',3'- 5x(2'OMe, 3'P(S))	CLTA4mg ON- target	<u>GsCsAsGsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAU AGCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUU <u>sUsUsUsU</u> (SEQ ID NO: 109)	113
257	CLTA4_5',3'- 5x(2'OMe, 3'P(S))	CLTA4mg OFF5- target	<u>GsCsAsGsAs</u> GAUGUAGUGUUUCCACAGUUUAAGAGCUAUGCUGGAAACAGCAU AGCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUU <u>sUsUsUsU</u> (SEQ ID NO: 109)	113
2'OMethyl,3'PhosphorothioPACE-modified sgRNA				
258	CLTA1_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA1 ON1- target	<u>A*sG*sU*</u> CCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCA UAGCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACC GAGUCGGUGCUUU <u>sU*sU*sU</u> (SEQ ID NO: 110)	113
259	CLTA1_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA1mg ON1- target	<u>A*sG*sU*</u> CCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCA UAGCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACC GAGUCGGUGCUUU <u>sU*sU*sU</u> (SEQ ID NO: 110)	113
260	CLTA1_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA1mg ON1- target	<u>A*sG*sU*</u> CCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCA UAGCAAGUUAAAAGGCUGGUACGGGUUAUCACUUGAAAAGUGGCACC GAGUCGGUGCUUU <u>sU*sU*sU</u> (SEQ ID NO: 110)	113

US 10,337,001 B2

97

98

TABLE 3-continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
261	CLTA1_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA1mg OFF1- target	A*sG*sU* sCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCA UAGCAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACC GAGUCGGGCUUUU <u>U*sU*sU</u> (SEQ ID NO: 110)	113
262	CLTA1_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA1mg OFF3- target	A*sG*sU* sCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCA UAGCAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACC GAGUCGGGCUUUU <u>U*sU*sU</u> (SEQ ID NO: 110)	113
263	CLTA1_5',3'- 1x(2'OMe, 3'thioPACE)	CLTA1mg ON1- target	A*sGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG UCGGGUGGUUUUU <u>*sU</u> (SEQ ID NO: 111)	113
264	CLTA1_5',3'- 1x(2'OMe, 3'thioPACE)	CLTA1mg ON1- target	A*sGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG UCGGGUGGUUUUU <u>*sU</u> (SEQ ID NO: 111)	113
265	CLTA1_5',3'- 1x(2'OMe, 3'thioPACE)	CLTA1mg OFF1- target	A*sGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG UCGGGUGGUUUUU <u>*sU</u> (SEQ ID NO: 111)	113
266	CLTA1_5',3'- 1x(2'OMe, 3'thioPACE)	CLTA1mg OFF3- target	A*sGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG UCGGGUGGUUUUU <u>*sU</u> (SEQ ID NO: 111)	113
267	CLTA1_5',3'-3x(2'OMe, 3'thioPACE)_75mer	CLTA1 ON1- target	A*sG*sU* sCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGC AUAGCAAGGUAAAAGGUAGCGUAGGU <u>C* sU*sU*sU</u> (SEQ ID NO: 112)	75
268	CLTA1_5',3'-1x(2'OMe, 3'thioPACE)_74mer	CLTA1mg ON1- target	A*sGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGGUAAAAGGUAGCGUAGGU <u>*sU</u> (SEQ ID NO: 112)	74
269	CLTA1_5',3'-1x(2'OMe, 3'thioPACE)_75mer	CLTA1mg ON1- target	A*sGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGGUAAAAGGUAGCGUAGGU <u>*sU</u> (SEQ ID NO: 112)	75
270	CLTA1_5',3'-1x(2'OMe, 3'thioPACE)_77mer	CLTA1mg ON1- target	A*sGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUAG CAAGGUAAAAGGUAGCGUAGGU <u>*sC</u> (SEQ ID NO: 113)	77
271	CLTA1_5',3'-1x(2'OMe, 3'thioPACE)_77mer	CLTA1mg ON1- target	G*sAGUCCUCAUCUCCCUCAGCGUUUAAGAGCUAUGCUGGUACAGCAUA GCAAGGUAAAAGGUAGCGUAGGU <u>AU*sC</u> (SEQ ID NO: 114)	78
272	CLTA4_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA4 ON- target	G*sC*sA* sGAUGUAGGUUUCACAGGUAAAAGAGCUAUGCUGGAAACAGCA UAGCAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACC GAGUCGGUGGUUUU <u>*sU*sU*sU</u> (SEQ ID NO: 115)	113
273	CLTA4_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA4 ON- target	G*sC*sA* sGAUGUAGGUUUCACAGGUAAAAGAGCUAUGCUGGAAACAGCA UAGCAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACC GAGUCGGUGGUUUU <u>*sU*sU*sU</u> (SEQ ID NO: 115)	113
274	CLTA4_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA4 ON- target	G*sC*sA* sGAUGUAGGUUUCACAGGUAAAAGAGCUAUGCUGGAAACAGCA UAGCAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACC GAGUCGGUGGUUUU <u>*sU*sU*sU</u> (SEQ ID NO: 115)	113
275	CLTA4_5',3'- 3x(2'OMe, 3'thioPACE)	CLTA4mg OFF5- target	G*sC*sA* sGAUGUAGGUUUCACAGGUAAAAGAGCUAUGCUGGAAACAGCA UAGCAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACC GAGUCGGUGGUUUU <u>*sU*sU*sU</u> (SEQ ID NO: 115)	113
276	CLTA4_5',3'- 1x(2'OMe, 3'thioPACE)	CLTA4mg ON- target	G*sCAGAUGUAGGUUUCACAGGUAAAAGAGCUAUGCUGGAAACAGCAUAG CAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG UCGGGUGGUUUUU <u>*sU</u> (SEQ ID NO: 116)	113
277	CLTA4_5',3'- 1x(2'OMe, 3'thioPACE)	CLTA4mg ON- target	G*sCAGAUGUAGGUUUCACAGGUAAAAGAGCUAUGCUGGAAACAGCAUAG CAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG UCGGGUGGUUUUU <u>*sU</u> (SEQ ID NO: 116)	113
278	CLTA4_5',3'- 1x(2'OMe, 3'thioPACE)	CLTA4mg ON- target	G*sCAGAUGUAGGUUUCACAGGUAAAAGAGCUAUGCUGGAAACAGCAUAG CAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG UCGGGUGGUUUUU <u>*sU</u> (SEQ ID NO: 116)	113
279	CLTA4_5',3'- 1x(2'OMe, 3'thioPACE)	CLTA4mg OFF5- target	G*sCAGAUGUAGGUUUCACAGGUAAAAGAGCUAUGCUGGAAACAGCAUAG CAAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG UCGGGUGGUUUUU <u>*sU</u> (SEQ ID NO: 116)	113
2-aminoA-modified sgRNA (including unmodified controls)				
280	EN1	EN1mg ON- target	GAUGUUGUGCUAGAAAAAGGUUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGGUAAAAGGUAGCGUAGGUUAUCACUUGAAAAGUGGCACCGAG CGGUGGUUUUU (SEQ ID NO: 117)	113

US 10,337,001 B2

99

100

TABLE 3 -continued

Entry #	Guide RNA Construct	Target DNA Construct	RNA sequence (5'→3')	RNA length
281 EN1		EN1mg OFF-target	GAUGUUUGUCGAUGAAAAGGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAAGGCUAGUCGUUAUCACUUGAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 117)	113
282 EN1_2aminoA + 16		EN1mg ON-target	GAUGUUUGUCGAUGAA (2aA) AAGGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAAGGCUAGUCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 118)	113
283 EN1_2aminoA + 16		EN1mg OFF-target	GAUGUUUGUCGAUGAA (2aA) AAGGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAAGGCUAGUCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 118)	113
284 PCDHA4		PCDHA4mg ON-target	GAUUUAGACGAAGGAUGUAAGGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAAGGCUAGUCGUUAUCACUUGAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 119)	113
285 PCDHA4		PCDHA4mg OFF-target	GAUUUAGACGAAGGAUGUAAGGUUUAAGAGCUAUGCUGGUACAGCAUAGC AAGUUUAAAAGGCUAGUCGUUAUCACUUGAAAAGUGGCACCGAGU CGGUGCUUUUUU (SEQ ID NO: 119)	113
286 PCDHA4_2aminoA + 15		PCDHA4mg ON-target	GAUUUAGACGAAGG (2aA) UUGAAGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAAGGCUAGUCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 120)	113
287 PCDHA4_2aminoA + 15		PCDHA4mg OFF-target	GAUUUAGACGAAGG (2aA) UUGAAGUUUAAGAGCUAUGCUGGUACAGCAU AGCAAGUUAAAAGGCUAGUCGUUAUCACUUGAAAAGUGGCACCG AGUCGGUGCUUUUUU (SEQ ID NO: 120)	113
5-methylU-modified sgRNA				
288 CLTA4_21x(5-MeU)		CLTA4mg ON-target	GCAGA (5mU) G (5mU) AG (5mU) G (5mU) (5mU) CCACAGUUUAAG AGC (5mU) A (5mU) GC (5mU) GG (5mU) AACAGCA (5mU) AGCAAGUUUA AUAAGGCUAGUCGUUAUCAAC (5mU) (5mU) GAAAAAG (5mU) GGCACCG AGUCGG (5mU) GC (5mU) (5mU) (5mU) (5mU) (5mU) U (SEQ ID NO: 121)	113
289 CLTA4_21x(5-MeU)		CLTA4mg OFF-5'-target	GCAGA (5mU) G (5mU) AG (5mU) G (5mU) (5mU) CCACAGUUUAAG AGC (5mU) A (5mU) GC (5mU) GG (5mU) AACAGCA (5mU) AGCAAGUUUA AUAAGGCUAGUCGUUAUCAAC (5mU) (5mU) GAAAAAG (5mU) GGCACCG AGUCGG (5mU) GC (5mU) (5mU) (5mU) (5mU) (5mU) U (SEQ ID NO: 121)	113
Z base-modified sgRNA				
290 CLTA1_ZZ_70,71		CLTA1 ON1-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAAGGCUAGZGUUAUCACUUGAAAAGUGGCACCGAGUC GGUGCUUUUUU (SEQ ID NO: 122)	113
291 CLTA1 ZZ_95,96		CLTA1 ON1-target	AGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AGUUUAAAAGGCUAGUCGUUAUCACUUGAAAAGUGGCACZZGAGUC GGUGCUUUUUU (SEQ ID NO: 122)	113
sgRNA modified to disfavor misfolding				
292 CLTA1_opti_short_5',3'-1x(2'OMe, 3'thioPACE)_2'OMe_54,57		CLTA1mg ON1-target	A*sGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAUAGCA AUAAGGUUA <u>UCCGUUA</u> ACAAGAAAUGUGGCACCGAGUCGGUCUU U*sU (SEQ ID NO: 123)	100
293 CLTA1_opti_short_5',3'-1x(2'OMe, 3'thioPACE)_2'OMe_64,67		CLTA1mg ON1-target	A*sGUCCUCAUCCCCUCAAGCGUUUAAGAGCUAUGCUGGUACAGCAU GCAAGUUAAAAGGUUA <u>UCCGUUA</u> ACAAGAAAUGUGGCACCGA GUCGGUGCUUUUU <u>*sU</u> (SEQ ID NO: 124)	113

N = 2'OMeN = 2'deoxyNs = 3'P(S)N* = 3'-PACEN*s = 3'-thioPACEN*s = 2'OMe,3'-PACENs = 2'OMe, 3'P(S)N_o = 5'-overhang (5' to the 20-nt guide sequence); "NC" means the overhang is not complementary to protospacer-adjacent sequence

(2sU) = 2-thioU

(2aA) = 2-aminoA

(5mU) = 5-methylU

Z = Z base

Intfl = Fluorophore incorporated at an internal position in the RNA sequence

US 10,337,001 B2

101

The DNA target constructs in Table 3 had the following sequences:

CLTA1 ON1-target:	AGAATTAACTGTGGTCACATTGCTTATCGACTGGCTCATCTCACAGCTCATC TTACGCAAGTTGATGAGTATGCCAGTCACCTTCATTTGTTGAATGTTCCCGTG ACATGCGAGTCTGCGACCAGTGTGCCCGGATTGAATTCTCAAGGGTGGTGATA GATGCTACGGTGGTATGCCAGTCAGTCCTCATCTCCCAAGGAGCCCC GCTGGTGGTCGGAGTCCTAGTGAAGGCACCAATATAAGTGGTCGTGCAAGCAAC TGCCACGCTCCACCCCTCGAGGTGCTAACATAACGTAAGGCACGAGTAAACA AGATCGATAGCAAGAACATGGTATAGACTGACGGAGAGCTGCCATTAGTCTGA (SEQ ID NO: 10)
CLTA1 OFF1-target:	AGAATTAACTGTGGTCACATTGCTTATCGACTGGCTCATCTCACAGCTCATC TTACGCAAGTTGATGAGTATGCCAGTCACCTTCATTTGTTGAATGTTCCCGTG ACATGCGAGTCTGCGACCAGTGTGCCCGGATTGAATTCTCAAGGGTGGTGATA GATGCTACGGTGGTATGCCAGTCAGTCCTCATCTCCCAAGGAGCCCC CCCTGGGGTGGAGTCCTAGTGAAGGCACCAATATAAGTGGTCGTGCAAGCAAC TGCCACGCTCCACCCCTCGAGGTGCTAACATAACGTAAGGCACGAGTAAACA AGATCGATAGCAAGAACATGGTATAGACTGACGGAGAGCTGCCATTAGTCTGA (SEQ ID NO: 11)
CLTA1 OFF2-target:	AGAATTAACTGTGGTCACATTGCTTATCGACTGGCTCATCTCACAGCTCATC TTACGCAAGTTGATGAGTATGCCAGTCACCTTCATTTGTTGAATGTTCCCGTG ACATGCGAGTCTGCGACCAGTGTGCCCGGATTGAATTCTCAAGGGTGGTGATA GATGCTACGGTGGTATGCCAGTCAGTCCTCATCTCCCAAGGAGCCCC ACTTTGGGTGGAGTCCTAGTGAAGGCACCAATATAAGTGGTCGTGCAAGCAAC TGCCACGCTCCACCCCTCGAGGTGCTAACATAACGTAAGGCACGAGTAAACA AGATCGATAGCAAGAACATGGTATAGACTGACGGAGAGCTGCCATTAGTCTGA (SEQ ID NO: 12)
CLTA1 OFF3-target:	AGAATTAACTGTGGTCACATTGCTTATCGACTGGCTCATCTCACAGCTCATC TTACGCAAGTTGATGAGTATGCCAGTCACCTTCATTTGTTGAATGTTCCCGTG ACATGCGAGTCTGCGACCAGTGTGCCCGGATTGAATTCTCAAGGGTGGTGATA GATGCTACGGTGGTATGCCAGTCAGTCCTCATCTCCCAAGGAGCCCC AGGCTGGGGTGGAGTCCTAGTGAAGGCACCAATATAAGTGGTCGTGCAAGCAAC TGCCACGCTCCACCCCTCGAGGTGCTAACATAACGTAAGGCACGAGTAAACA AGATCGATAGCAAGAACATGGTATAGACTGACGGAGAGCTGCCATTAGTCTGA (SEQ ID NO: 13)
CLTA1mg ON1-target:	GCGTTTCTGGGTGAGCAAAACAGGAAGGCAAATGCCGAAAAAGGAATAAGG GCGACACGGAAATGTGAATACTCATACTCTTCCTTTCAATATTATTGAAGCAT TTATCAGGGTTATTGTCATGAGCGATACATATTGAATGTTAGAAAAATA AACAAATAGGGGTTCCCGCAGTTCAGCTTCCAGCCAAAGTGCACCTAAATTGTAAGCG TTAATATTGTTAAATTCCCGTTAAATTGTTAAATCAGCTCATTTTAAAC CAATAGGCCAAATCGGAAATCCCTATAAATCAAAGAATAGACCGAGATAGG GTTAGTGTGTTCCAGTTGGAAACAAGAGTCCACTATAAGAACGTGGACTCCA ACGTCAAAGGGCAACCGCTATCAGGGCAGTGGCCACTACGTGAACCATCA CCTAATCAAGTTTGGGTGAGGTGCGTAAGCAGTAAATCGGAACCTAA AGGGAGCCCCGATTAGAGCTTGACGGGAAAGCGCGAACGTGGCGAGAAAGG AAGGAAGAAAAGCAGGGAGCGGGCCTAGGGCCTGGCAAGTGTAGCGGTACCG CTGCGCGTAACACACCACCCCGCGCTTAATGCGCCCTACAGGGCGCTCCCA TTCGCGATTAGGGCTCGCAACTGTGGGAAGGGCATCGGTGCGGGCTTTCGCG TATTACGCCAGTGGCAAAGGGGATGTGCTGCAAGGCATAAGTGGTAACG CCAGGGTTTCCCAGTCAGCACGTTGAAACGACGCCAGTGTAGCGCGCTAAAT CGACTCATAGGGCAATTGGGTACGATGCGCCCTCGCAGGCCAAAGATG TCTCCCGCATGCCCTAGTCCTCATCTCCCAAGCAGGCCCTGCTGGTCACTGA AGAGCCACCTGTGCGCGTATGCGCTCCAGCTTGTCCCTTAGTGAGGG TTAATTGCGCCTTGGCTAAATCATGTCATAGCTGTTCTGTGAAATTGTTA TCCGCTCACACACACACACACAGGCGGAAAGCATAAAAGTGTAAAGCCTGGG GTGCGCTAATAGGTGAGCTCACTCACATTAATTGCGTGGCCTACTGCCCGCTTTC CAGTCGGAAACCTGTGCGCAGTCGCTTAATGAATCGCCACGCCGGGGAG AGGCCTGGTGTGCTATTGGGCTCTCCGCTCTCGCTACTGACTCGTGTGCGCT CGGTCTGGCTGGCGTACGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT TCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAGGCCAGCAAA AGGCCAGGAAACCGTAAAAGGCCGCGTTGCTGGCTTTCCATAGGCTCCGCCCTA CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCAACCCGACAGGA CTAGTAAAGATAACAGGGCTTCCCCCTGGAAAGCACGACTTACCGCTCTGGCC GACCCCTGCCCTACGGGATACCTGTGCCCTTCTCCCTCGGGAAAGCGTGGGCC TTTCTCATAGCTCACCGCTGAGGTATCTCGAGTCGTTGAGGTGCTCGTCCAAAG CTGGGCTGTGTCAGCAACCCCGTTCAGCCGACCGCTCGCCTTATCGGTAA ATCTCGTCTTGAGTCACCCGGTAAGACACGACTTATCGCCTACTGGCAGCAGCA CTGGTAACAGGGATTAGCAGAGCGAGGTATGAGGTGCTACAGAGTTCTTGAGG TGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTGTTAGTGTGCTCGT GAAGGCCAGTACCTTCGAAAAAGAGTGGTAGCTTGTGATCCGGAAACAAACCA CGCCTGGTAGCGGTGTTTTGGTTGCAAGCAGCAGATTACGCCAGGAGAAAAAAA GGATCTCAAGAAGATCTTGATTTCTACGGGGTCTGAGCTCAGTGGAAACGA AAACTCAGGTTAAGGGATTGGTATGAGGATTATCAAAAGGATCTCACCTAGA TCCTTTAAATTAAAATGAAGTTAAATCAATCAAAGTATATAGTGTAAACT TGGTGTGACAGTTACCAATGCTTAATCAGTGAGGACACTATCTCAGCGATCTGT ATTTCGTTCATCCATAGTTGCGTGAACCTCCCGTGTGAGATAACTACGATACGGG AGGGCTTACCATCTGGCCCCAGTGTGCAATGATAACCGCGAGACCCACGCTCACCG

102

US 10,337,001 B2

103

104

- continued

GCTCCAGATTATCAGCAATAAACCGAGGCCAGCGGAAGGGCCAGCGCAGAAGTGG
 TCCCTGCAACTTATCGGCCTCATCCAGTCTATTAAATTGTGCGGGAAAGCTAGAG
 TAAGTAGTTGCCAGTTAATAGTTGCGCAACGTTGTGCCATTGCTACAGGCATC
 GTGGTGTACGCTCGTCGTTGGTATGGCTCATTAGCCTCCGGTCCAAACGATC
 AAGGGAGTTACATGTCGTTGGTATGGCTCATTAGCCTCCGGTCCAAACGATC
 CTCCGATCGTGTGAGTAAGTGGCCGAGTGTATCACTCATGGTTATGGCA
 GCACTGCATAATTCTTACTGTATGCCCATCGTAAGATGCTTTCTGTGACTGG
 TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCCGACCGAGTTGCTCTT
 GCCCGCGTCATAACGGATAATACCGCGCACATAGC (SEQ ID NO: 14)

CLTA1mg OFF1-target:

GCCTTCTGGGTGAGAAAAACAGGAAGGCAAATGCCGAAAAAGGAAATAAGG
 GCGACACGGAAATGTGAATACTCATACTCTTCTTTCAATTATTAGCAT
 TTATCAGGGTTATGTCATGAGCGGATACATATTAGTATTTAGAAAATA
 AACAAATAGGGTTCCGCACATTCCCCAAAAGTGCACCTAAATTGTAACCG
 TTAATTGTTGTTAAATTCCGTTAAATTGTTGTTAAATCAGCTATTGTTAAC
 CAATAGGCCAAATCGGAAATCCCTATAAAATCAAAGAATAGACCGAGATAGG
 GTTGAGTGTGTTCCAGTTGGAAAGAGTCCACTATTAAAGAACGTGGAACCTCA
 ACGTCAAAGGGAAACCGTCTATCAGGGCGATGCCACTACGTGAACCATCA
 CCCTAAATCAAGTTTTGGGTCGAGGTGCGTAAAGCATAATCGGAACCCCAA
 AGGGAGCCCCGATTAGAGCTTGACGGGAAAGCGCGAACGTGGCGAGAAAAGG
 AAGGGAGAAAGCAGGGAGCGGGCTAAGGGCGCTGGCAAGTGTAGCGGTACCG
 CTGCGGTAAACACCACCCGCCGCGCTTAATGCGCGTACAGGGCGCGTCCCA
 TTCGCGATTAGGCTCGCAACTGTTGGGAGGGCGATCGGTGCGGGCTTTCG
 TATTACGCCAGTGGCAAAGGGGATGTGCTGCAAGGGATAAGTGGTAACG
 CCAGGGTTTCCCAGTCAGACGTTGTAACACGACGCCAGTGGCGAGCGTAA
 CGACTCATAGGGCAATTGGGCTGAGGTGCGTAAAGCATAATCGGAACCCCAA
 AGGGAGGGAAACCGTCTAGCTGCGTAACTCCGCTTGTGACTGAGGAGG
 TCTCTGTATGCACTGCTCTCACTCCCTCAAGCAGGGCACCTTGGTGCACCTG
 ACAAAACCGCTCTGCCGTGATATGCGACTCCAGCTTGTGCTTGTGAGGG
 TTAATTGCGCCTTGGCTAATCATGTCATAGCTGTTCTGTGAAATTGTTA
 TCCGCTCACAAATCCACACACATACGGAGCGGAAGCATAAAGTGTAAAGGCTGG
 GTGCGTAAATGAGTGAACCTAACATTAATTGCGTTGCGCTACTGCCGTTTC
 CAGTCGGGAAACCTGCGTCCAGTGCCTTAATGAATCGGCCACGCCGGGGAG
 AGGCCTGTTGCGTATTGGGCGCTTCCGCTTCTGCGTCACTGACTCGTGC
 CGGCGTACGGCGTGTGAGTGTGCGTCACTCAGCTCAGCTTGTGCGTCTG
 TCCACAGAATCAGGGATAACGGAGAACACATGTGAGCAGGGAGCAGCAA
 GGCCAGGAAACCGTAAAGGCCGCGTTGCTGGCGTTTCCATAGGCTCCGCC
 CTGACGGCATCACAAATCGACGCTCAAGTCAGGGTGGCAACCCGACAGGA
 CTATAAAGATAACGGCGTACCTCCCTGGAGCTCCCTCGTGCCTCTCTGTT
 GACCCCTGCCGTTACGGGATACCTGCGCCTTCTCCCTCGGGAGGCTGGGC
 TTTCTCATAGCTCACGCTGTAGGTATCTCGTGTAGGTGCTCGTCCAAAG
 CTGGGCTGTTGTCACGAAACCCCCGTTGACGCCGACCGCTGCGCTTATCGGTA
 AATCTGCTTGTGAGCTTACGGGCTGAGACACGACTTATCGCCACTGGCAGCAGCA
 CTGGTAACAGGATTAGCAGAGCGAGGTATGAGGCGTGTACAGAGTTCTG
 TGTTGGCTAAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGTCTG
 GAAGCCAGTTACCTGGAAAAGAGTGTGAGCTTGTGAGCTTGTGAGGCA
 ACCGCTGTTAGCGGTGTTTGTGAGCAGCAGGAGATTACGGGAGAAAAAAA
 GGATCTCAAGAAGATCTTGTATTTCTACGGGCTGAGCTGAGCTGAGGAA
 AAACACTACGTTAAGGGATTGGTATGAGATTATCAAAGGATCTCACCTAGA
 TCCTTTAAATTAAAATGAAGTTAAATCAATCTAAAGTATATGAGTAAACT
 TGGCTGAGCTTACCAATCTAAATGAGGAGCAGTCTGAGCTGAGCTG
 ATTTCGTTCTACATCATGGCTGACTCCCGTGTGAGATAACTACGATACGG
 AGGGCTTACCATCTGCCCGTGTGCAATGATACCGAGACCCACGCTCACCG
 GCTCCAGATTATCAGCAATAAACCCAGGCCAGCGGAAGGGCGAGCGCAGAGTGG
 TCCGCAACTTATCGCCCTCATCCAGTCTATTAAATTGCGGCTGAGCTGAG
 TAAGTAGTGTGCGTAAAGTGGCAACGGTGTGAGCTTGTGAGGAGCTGAG
 GTGGCTGACGCTCGTGTGGTATGGCTTACAGCTCGGTTCCAAACGATC
 AAGGGAGGTTACATGATCCCCCATGGTGTGCAAAAGCGGTTAGCTCCTCG
 CTGGCATCGTGTGAGAAGTGGCCGAGTGTATCACTCATGGTTATGGCA
 GCACTGATAATTCTTACTGTATGCCCATCGTAAGATGCTTTCTGTGACTGG
 TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCCGACCGAGTTGCTCTT
 GCCCGCGTCATAACGGATAATACCGCGCACATAGC (SEQ ID NO: 15)

CLTA1mg _OFF3-target:

GCCTTCTGGGTGAGAAAAACAGGAAGGCAAATGCCGAAAAAGGAAATAAGG
 GCGACACGGAAATGTGAATACTCATACTCTTCTTTCAATTATTAGCAT
 TTATCAGGGTTATGTCATGAGCGGATACATATTAGTATTTAGAAAATA
 AACAAATAGGGTTCCGCACATTCCCCAAAAGTGCACCTAAATTGTAACCG
 TTAATTGTTGTTAAATTGCGTTAAATTGCGTTAAATCAGCTCATTGTTAAC
 CAATAGGCCAAATCGGAAATCCCTATAAAATCAAAGAATAGACCGAGATAGG
 GTTGAGTGTGTTCCAGTTGGAAAGAGTCCACTATTAAAGAACGTGGAACCTCA
 ACGTCAAAGGGAAACCGTCTATCAGGGCGATGCCACTACGTGAACCATCA
 CCCTAAATCAAGTTTTGGGTCGAGGTGCGTAAAGCAGCTAAATCGGAACCCCAA
 AGGGAGCCCCGATTAGAGCTTGACGGGAAAGCGCGAACGTGGCGAGAAAAGG
 AAGGGAGAAAGCAGGGAGAGCGGGCGCTAGGGCGTGGCAAGTGTAGCGGTACCG
 CTGCGGTAAACCCACACCCGCCGCGCTTAATGCCGCTACAGGGCGCGTCCCA
 TTCGCGATTAGGCTGCCACTGTTGGGAGGGCGATCGGTGCGGGCTTTCG
 TATTACGCCAGCTGGGAAAGGGGATGTGCTGCAAGGGATTAAAGTGGTAACG
 CCAGGGTTTCCCAGTCAGACGTTGTAACACGACGCCAGTGGCGCGTAAATA
 CGACTCACTAGGGCAATTGGGTACGATCGATGCGGCCTCAGGAGAGGGAGGCC
 TGCTCATCTCCAGGCCACTCCCTCATCCCCCTCAAGCCGGCTGAGGAGGCT

US 10,337,001 B2

105

- continued

106

AAAGCTGTCTTGCGCGTGTATGCGACTCCAGCTTGTGCCCTTAGTGAGGG
 TTAATTGCGCGTTGGCGTAATCATGTCATAGCTTCTGTGTGAAATTGTTA
 TCGCCTACAATTCCACACACATAGCGAGCGGAAGCATAAGTGTAAAGCCTGGG
 GTGCCTAATGAGTGAGCTAACACATCACATTAAATGCGTGCCTCACTGCCCGCTTCA
 CAGTCGGAAACCTGTCTGTGCCAGCTGCAATTAAATCGGCCAACCGCGGGGAG
 AGGCCTTGTGCTATTGGCGCTCTCCGCTTCCCTCGCTCACTGACTCGTGCCT
 CGGTCTCGCTGCCGGAGCGGTACAGCTCACTCAAAGCGGTAAATCGGTTA
 TCCACAGAATCAGGGATAACGCAGGAAGAACATGTGAGCAGAAAAGGCCAGCAAA
 GGCCAGGAACCGTAAAAGGCCGCTTGTCTGGCGTTTCCATAGGCTCGCCCCCCC
 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAACCCGACAGGA
 CTATAAAGATAACCGCGTTTCCCTGGAGGCTCCCTCGCTCTCGTCTCGTCC
 GACCCCTGCCCTACCGATAACCTGTCCGCTTCTCCCTCGGGAAAGCGTGGCGC
 TTCTCATAGCTACCGCTGTAGGTATCTCAGTCGGTGTAGGTCTCGTCCAAAG
 CTGGCTGTGTCAGGAAACCCCGCTTCAGCCGACCCGCTCGCCTTATCCGTA
 CTATCGCTTGTGAGTCAACCCGGTAAGAACACGACTTACGCCACTGGCAGCAGCA
 CTGGTAACAGGATTAGCAGAGCAGGGTATGAGGCGGTGTCAGAGGTTCTGAAG
 TGGTGGCCCTAACCTGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCT
 GAAGCCAGTTACCTTCGGAAAAAGAGTGTGGTAGGTCTTGATCCGGAAACAAACCA
 CCGCTGGTAGCGGTGTTTTTTGTTGCAAGCAGCAGATTACCGCAGAAAAAAA
 GGATCTCAAGAAGATCCTTGATCTTCTACGGGCTGACGCTAGTGGAACGA
 AAACCTCACTGTTAGGGATTGGTATGAGATTATCAAAAAGGATCTTACCTAGA
 TCCCTTAAATTAAATGAGTTTAAATCAATCTAAAGTATATATGAGTAAACT
 TGGTCTGACAGTTACCAATGTTAAATCAGTGAGGCACTTACAGCAGATCTGCT
 ATTTCTGTCATCCATAGTTGCTGACTCCCGTGTAGATAACTACGATACGGG
 AGGGCTTACCATCTGCCCTGGAGCTGCTGCAATGATAACCGCAGACCCACGCTCACCG
 GCTCCAGATTATCAGCAATAACCCAGCAGCGGAAAGGGCGAGCGCAGAAGTGG
 TCCCTGCAACTTATCCGCTCATCCAGCTTAAATTGTCAGGTTGCTGCACTACGGCATC
 TAAGTAGTTGCCAGTTAATAGTTGCGCAACGTTGTGCTGCACTACGGCATC
 GTGGTGTACCGCTCGTGTGGTATGGCTTCATTAGCTCCGGTTCCAACGATC
 AAGGGAGTTACATGATCCCCCATGTTGCAAAAAAGCGGTTAGCTCCCGT
 CTCCGATCTGTTGCAAGAGTAAGTTGCCAGTGTATCACTCATGGTTATGGCA
 GCACTGCAATTCTCTACTGTCTGATGCCATCCGTAAGATGCTTTCTGTGACTGG
 TGAGTACTCAACCAAGTCACTTGAGAATAGTGTATGCCGACCGAGTTGCTCTT
 GCCCGCGTCATAACGGATAATACCGCGCACATAGC (SEQ ID NO: 16)

CLTA4 ON-
target:

GCGGGCAAGAGCAACTCGTCGCCGACATACACTATTCAGAATGACTGGTTGA
 GTACTCACCAGTCACAGAAAAGCATCTACCGATGGCATGACAGTAAGAGAAATTAT
 GCAGTGTGCCATACCATGAGTGTATAACACTGCCGCAACTTACTCTGACAACG
 ATCGGAGGACCGAAGGAGCTAACCGCTTTTGCAACACATGGGGATCATGTAAC
 TCGCTTGATCGTGGAAACCGAGCTGAATGAAGCCATACCAACGACGAGCGTG
 ACACCAAGATGCCCTGAGCAATGGAACAAAGCTGGCAGGAAACTTAACTGGAA
 CTACTTACTCTAGCTCCGGCAACAAATTAAATAGACTGGATGGAGGGGGATAAATG
 TGCAGGACCACTTCTGCCCTCGGCTTCCGGCTGGCTTATTGCTGATAAAAT
 CTGGAGCCGGTGAGCGTGGGTCGCGGTATCATTGCACTGGGGAGATGGT
 AAGCCCTCCGTATCTAGTTATCTACACGACGGGAGTCAGGAACATATGGATGA
 AGCAAAATAGACAGATGCTGAGATAGGTGCTCACTGATTAAGCATTGTAACCTG
 CAGCCAAGTACTCATATAACTTGTGATTTAAACTTCTTAAATTTAAATTT
 AAAAGGATCTAGGTGAAGATCTTTTGATAATCTCATGCAAAATCCCTAACG
 TGAGTTTCGTTCCactGaGCGTCAGACCCGTAAGAAAAGATCAAAGGATCTTCT
 GAGATCTTTTCTGCGGTAATCTGCTGCTTCCAAACAAAACCCACCGCTA
 CCAGCGGTTGGTTGGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC
 TGGCTCAGCAGGCCAGATACAAATACTGTTCTCTAGTGTAGCCGTAGTTAG
 GCCACCACTCAAGAACCTGTAGCACCGCCACATACCTCGCTCTGCTAATCCTGT
 TACCACTGGCTCTGGCAGTGGCGATAAGTGTGCTTACCGGGTTGGACTCAAG
 ACGATAGTACCGGATAACGGCGCAGCGCTGGGCTGAACGGGGTTCTGCAACAC
 AGCCACGTTGGAGCGAACGACCTACACCGAACCTACAGCTGAGGATACTACAGCGTGAGCTA
 TGAGAAAGGCCACGCTCCGAAGGGAGAAGGCCAGGGTATCCGGTAAGCGG
 CAGGGTCCGAACAGGGAGCGAGCGAGCTCCAGGGGAAACGCCCTGGTATC
 TTTATAGTCTGCTGGGTTGCCACCTCTGACTTGTGAGGCGCTGATTTTGATGC
 TCGTCAGGGGGCGGAGCCATGGAAAACCCAGCAACCGGGCTTTCAGGTT
 CCTGGCTTTGCTGCCCTTGTGTCACATGTTCTCTGCTTATCCCTGATT
 CTGTGGATAACGGTATACCCCTTGAGTGAGCTGATACCGCTGCCAGCGGA
 ACGACCGAGCGCAGCGAGCTGAGCGAGGAGCGGAAACGCCCTGGTATC
 ACCGCCCTCTGCCCGCGCTGGCGATTCAATTGCACTGAGGCGCTGAGCGAACAGGTTTC
 CCGACTGGAAGCGGGAGCTGAGCGCAACGCAATTATGTGAGTTAGCTACTCAT
 TAGGACCCAGGTTACCTTATGCTTCCGGCTCGTATGTTGTGGAATTGT
 GAGCGGATAACAAATTACACAGGAACAGCTATGACATGATTACGCCAACGCG
 CAATTAAACCTCATGAGGAAACAAAGCTGGTACCGGGCCCCCTCGACACC
 AGTTGCAATTGCTGTTGTAATTGCTCAATTCTGAGGCCGGGAGCTGGC
 AGATGTAGTGTGTTCCACAGGGATCCTAGTTCTAGAGCGGCCACCGCGGTG
 GAGCTCAATTGCGCTATGAGTCGTTACCGGCCGCTCACTGGGGCTCGTT
 TACAACGCTGACTGGAAAACCTGGCTTACCAACTTAATGCCCTGCGAGCA
 CATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGACCGATGCCCTTC
 CCAACAGTTGCCAGCTGAAATGGGAATTGTAAGCGTTAATTTTGT
 AAAATTGCGTTAATTGTTGAAATGAGCTCATTTTAACCAATAGGCCAAA
 TCGGAAACATCCCTATAAACTAAAGAATAGGCCAGATAGGGTTGAGTGTGTT
 CCAGTTGGAAACAGAGTCCACTTAAAGAACGCTGACTCAACGTCAGGCTAAAGGG
 AAAACCGTCTATCAGGGCGATGCCCACTACGCTGAAACCATCACCTAAATCAAGT
 TTTGGGTCAGGGTCCGTAAGCAGTAAACGCACTGGGAGCCCCCGA

US 10,337,001 B2

107

- continued

108

TTTAGAGCTTGACGGGAAAGCCGCAACGTGGCGAGAAAAGGAAGGGAAAGAAGC
 GAAAGGAGCGGGCGTAGGGCGTGGCAAGTGTAGCGGTACGCGCTAACCA
 CCACACCGCGCGCTTAATGCCGCACAGGGCGTCAGGTGGACTTTCGG
 GGAAATGTGCGGAAACCCATTGTTTCTAAATACATTCAAATATGTA
 TCGCTCATGAGAACATAACCCGTATAATGCTCAATAATATTGAAAAGGAAGA
 GTATGAGTATTCAACATTCGGTGCCTTATTCCCTTTGCGGCATTTC
 CTTCTGTTTGTCAACCCAGAACGCTGGTGAAGTAAAGATGCTGAAGATCA
 GTTGGGTGCAAGAGTGGTTACATGCACTGGATCTAACAGCGGTAAGCTTG
 AGAGTTTCGCCAGAAGAAGCTTTCATGATGAGCACTTTAAAGTCTGCTA
 TGTGGCGGGTATTATCCGTATTGAC (SEQ ID NO: 17)

CLTA4mg ON-target: GCGTTCTGGGTGAGAAAAACAGGAAGGCAAATGCCGAAAAAGGAATAAGG
 GCGACACGGAATGTGAATACTCATACTCTCCCTTTCAATATTATTGAGCAT
 TTATCAGGGTATTGTCATGAGCGATACATATTGATGTTAGAAAATA
 AACAAATAGGGGTCGGCACATTCCCGAAAAGTGCACCTAAATTGTAAGGG
 TTAATATTGTTAAATTGCGTTAAATTGTTGTTAAATCAGCTATTTC
 CAATAGGCCAAATCCCTATAAATCAAAGAATAGACCGAGATAGG
 GTTGAGTGTTCAGTTGGAAACAAGGCTTCACTTTAAAGAACGTTGACTCCA
 ACGTCAAAGGCCAAAACCGCTATCAGGGCGATGCCACTACGTGAACCCTA
 CCTAATCAAGTTGGGTGAGGTGCGTAAAGCACTAAATCGGAACCTAA
 AGGGAGCCCGGTTAGAGCTGAGGGAAAGCCGGAACGTGGCGAGAAAGG
 AAGGGAAGAAGCGAAAGGAGCGCTAGGGCGCTGGCAAGTGTAGCGGTAC
 CTGCGCTAACCCACACCCGCCGCTTAATGCCGCTACAGGGCGCTCCCA
 TTCGCCCCATTAGGCTGCCACTGTGGGAAGGGATCGGTGCGGGCTTC
 TATTACGCCACTGGGAAAGGGGATGTGCTGAGGCTAAAGTGGTAACG
 CCAGGGTTTCCCAGTCAGCAGCTGCAACGCGAACAGCTGGCGTAAATA
 CGACTCACTAAGGGGAAATTGGGTACGATCGATGGGCTCAAGAGCTTCA
 GTAGGATAAGATATTGAGATGTAGTGTTCACAGGGTGGCTCTCAGTGC
 ACCGGAACCTGCTGCCGTATGAGCTCCAGCTTTGTTCCCTTACTGAGGG
 TTAATTGGCCTGGGTAATCATGGTCAAGCTGTGTTCTGTGAAATTGTTA
 TCCGGCTCACAACTCCACACATCGAGCGGAAGCATAAGGTAAAGCCTGG
 GTGCCCTAATGAGTGAGCTAACATCACATAATTGCGTGCCTACTGCCCTT
 CAGTCGGAAACCTGCTGCGCCAGCTGCAATTAAATGAACTGCCAACGCC
 AGGGGTTTGGGCTTGGGCTTGGGAAAGGCTCCCTGCTCAGTGCCT
 CTAGGGAACCTGCTGCCGTATGAGCTCCAGCTTTGTTCCCTTACTGAGGG
 TTAATTGGCCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG
 TCAAGGGGAAACGGTAAAGGCGCGCTGGCTGGCTGGCTGGCTTGGGCT
 CTAGTCGGAAACCCGGTAAAGACAGCTTATCGCACTGGCAGCAGCA
 CTGGTAAAGGATTAGCAGAGCAGGGTATGAGGCGGTGCTACAGAGTTCT
 GAGGGCTTAAACTACGGCTACACTAGAAGGACAGTTGGTATCTGCGCT
 GAAGGCAAGTGGCTGGGAAAGAGTGGTAGCTGTGATCGGCAAACAAACCA
 CGCCTGGTAGCGGGTTTTGGGCTGGCTGGCTGGCTGGCTGGCTGGCT
 GGATCTCAAGAAAGATCTTGTATCTTCTACGGGCTGACGCTGAGGCA
 AAACTCAGGTTAAGGGATTGGTACGAGATTATCAAAGGATCTCACCTAGA
 TCTTCTTAAATAAAATGAAAGTTAAATCAATCTAAAGTATATGAGTAAAC
 TGTTCTGAGTACCTAACATGTTAATCAGTGGGACTATCTCAGCGATCTG
 ATTTCTCATCCATAGTTGCTGACTCCCGTCGTGAGATAACTACGATACGG
 AGGGCTTACCATCTGCCCGCAGTGTGCAATGATACCGCAGACCCACGCT
 GCTCAGGATTATCACGAAACCCAGCGAGCCGAGGGCAGCAGAGTGG
 TCCGCACTTATGCCCTCATCCAGTCTTAAATTGTCGGGAAAGCTAGAG
 TAAGTAGGTTGCCAGTTAATGAGTTGCGCAACGTTGTGCTTGTGAG
 GTGGTGTACGCTCGCTGGTATGGCTTACGCTCCGGTCCAAACGATC
 AAGGGAGTTACATGCCCTCATGGTCAAGGAGGGTTAGCTCCCTTGGCTC
 CTCCGATCTGGTCAAGAGTAAGTGGGCGAGTGTATCACTCATGGTTATGG
 GCACTGCAATACTCTTACTGTATGCCCATCGTAAGATGCTTTCTGTGACTGG
 TGAGTACTCAACCAAGTCACTTGAGAATAGTGTAGCGGGGACCGAGTTGCT
 GCCCGGGCTAACACGGATAATACCGCGCACATAGC (SEQ ID NO: 18)

CLTA4mg OFF5-target: GCGTTCTGGGTGAGAAAAACAGGAAGGCAAATGCCGAAAAAGGAATAAGG
 GCGACACGGAATGTGAATACTCATACTCTCCCTTTCAATATTATTGAGCAT
 TTATCAGGGTATTGTCATGAGCGATACATATTGATGTTAGAAAATA
 AACAAATAGGGGTCGGCACATTCCCGAAAAGTGCACCTAAATTGTAAGGG
 TTAATATTGTTAAATTGCGTTAAATTGTTAAATCAGCTATTTC
 CAATAGGCCAAATCCCTATAAATCAAAGAATAGACCGAGATAGG
 GTTGAGTGTTCAGTTGGAAACAAGAGTCCACTTTAAAGAACGTTGACTCCA
 ACGTCAAAGGCCAAAACCGCTATCAGGGCGATGCCACTACGTGAACCATCA
 CCTAATCAAGTTGGGTGAGGTGCGTAAAGCATAATCGGAACCTAA
 AGGGAGCCCGATTAGAGCTTGAAGGGAAAGCCGCGAACGTGGCGAGAAAGG
 AAGGGAAGAAGCGAAAGGAGCGGGCTAGGGCGCTGGCAAGTGTAGCGGTAC
 CTGCGCTAACACCACACCCGCCGCTTAATGCCGCTACAGGGCGCTCCCA
 TTCGCCCCATTAGGCTGCCACTGTGGGAAGGGGATGTGCTGAAGGCGATTAAGTGGGTAACG
 CCAGGGTTTCCCAGTCAGCAGCTGAAACGAGCGCTAGGCGCAGTGC
 CGCGTAATA (SEQ ID NO: 19)

US 10,337,001 B2

109

110

- continued

GGAAGAGATGCATACAAACATATGAGTATTCCACAGGGAATACAATGGACAATG
 ACCTCAAGAGCAGGCCGTATATGCAGCTCCAGCTTGTCCCTTAGTGAGGG
 TTAATTGCGCCTTGGCTAATCATGGTCATAGCTGTTCTGTGAAATTGTTA
 TCCGCTACAATTCCACACAACATAAGCGAGCCGAAGCATAAGTGAAAGCCTGGG
 GTGCCTAATGAGTGAAGCTAACATCACCATTAACTGCCTGCGCTACTGCCCGCTTTC
 CAGTCGGAAACTCTGCGTGCAGCTGCAATTAAATGAATCGGCCAACGCGCGGGAG
 AGGCCTTGGCTATTGGGCCTTCTCCGCTCTCGCTACTGACTCGCTGCGT
 CGGTGCTTGGCTGCGGAGCGGTATCAGCTCACTCAAAGGCGGTATACGGTTA
 TCCACAGAATCAGGGATAACGCAGGAAAGAACATGAGCAGAAAAGGCCAGCAGAAA
 GGCGAGAACCGTAAAAAGGCCGCGTGGCTGGCTTTCCATAGGCTCCGCCCC
 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAACCCGACAGGA
 CTATAAAGATAACCAGGGTTTCCCCCTGGAAGGCTCCCTGCGCTCTCTGTTCC
 GACCCCTGCCCTACCGGATACCTGCGCTTCTCCCTTGGGAAGCGTGGCGC
 TTTCTCATAGCTACCGCTGAGGTTCTCAGTGGCTGAGGTCTTCGCTCCAAG
 CTGGCTGTTGCAACGGGAAACCCCGTCAAGCCGACCGCTGCGCTTATCGGTAA
 CTATCGTCTGAGTCAACCCCGTAAGACAGACTATCGCCTGCGCAGCAGCCA
 CTGGTAACAGGATTAGCAGAGCGAGGTATGAGGCTGCTACAGAGTTCTTGAAG
 TGGTGGCCTAACCGCTACACTAGAAGGAGCAGTTGGTATCTGCGCTCTGCT
 GAAGCCAGTAACTTGGAAAGAGTGGTAGCTTGTGATCCGCAAACAAACCA
 CCGCTGGTAGCGGTGGTTTTGGCAAGCAGCAGATTACGCGCAGAAAAAAA
 GGATCTCAAGAAGATCTTGTGATCTTCTACGGGCTCTGACGCTCAGTGGAAACGA
 AAACCTCACAGGTTTGGGATCTGGCTCATGAGATTCTAACAAAGGATCTCACCTAGA
 TCCCTTAAATAAAATGAAGTTTAAATCAATCTAAAGTATATGACTAAACT
 TGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATCTAGCGATCTGCT
 ATTTCTCATCCATAGTTGCTGACTCCCGTGTAGATAACTACGATACGGG
 AGGGCTTACCATCGGCCAGTGTGCAATGATACCGGAGACCCACGCTCACCG
 GCTCCAGATTATCACAAATAACCCAGGCCAGCGGAAGGGCCGAGCGCAGAAGTGG
 TCCGCAACTTTATCGCCCTCATCCAGTCTATTAAATGTTGCGGGAAAGCTAGAG
 TAAGTAGTTGCCAGTTAATAGTTGCGCAACGTTGCTGCAATTGCTACAGGCATC
 GTGGTGTACGCTGCTGTTGCTATCGCTCCAGTCCGGGTTCCAACAGATC
 AAGGGAGGTTACATGATCCCCCATGTTGCAAAAAGCGGTTAGCTCTCGGT
 CTCCGATCGTGTCAAGAAGTGTGGCCAGTGTATCACTCATGGTTATGGCA
 GCACGTGATAATTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGACTGG
 TGAGTACTGATAACCGGATAATACGGCCACATAGC (SEQ ID NO: 19)

IL2RGmg_ON
 target :
 GCGTTCTGGGTGAGCAAAAAGGAAGGCAAAATGCGCAAAAAGGGATAAGG
 GCGACACGGAAATGTAATCTACATCTTCCCTTCAATATTATGAAAGCAT
 TTATCAGGGTATTGTCATGAGGGATACATATTGAATGTTAGAAAATA
 AACAAATAGGGGTTCCGCGCACATTCCCAGAAAGTGGCACCTAAATTGTAAGCG
 TTAATTTGTTAAATTCGCTTAAATTGTTGTTAAATCAGCTATTGTTAAC
 CAATAGGGCAAAATGGCAAATCCCTATAAATCAGGAAATAGACCGAGATAGG
 GTTGAAGTGTGTTCCAGTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCA
 ACGTCAAAGGCCAAAAACCGTCTACAGGCGATGCCACTACGTGACCATCA
 CCCTAATCAAGTTGGGCTGAGGTGCGTAAAGCAGTAAATGGAACCTAA
 AGGGAGCCCCGATTAGAGCTTGAGGGAAAGCCGGAACCGTGGAGAAGG
 AAGGGAGGAAAGCGAAGGGAGGGCGCTAGGGCGTGGCAAGTGTAGCGTACG
 CTGCGCTAACACACCCGCCGCTTAATGCCGCTACAGGGCGCTCCCA
 TTCGCAATTAGGCTGCCACTGTGGGAAGGGCATCGTGCGGGCTCTTCGC
 TATTAGGCTGCCACTGGGATGTGCTGCAAGGGGATTAAAGTTGGGTAACG
 CCAGGGTTTCCAGTCACGCTGTTAAACGACGCCAGTGAGCCGCGTAATA
 CGACTACTATAGGGGAATTGGGATCGATCGATGCCCTCGGGCAGCTGCAGGA
 ATAAGAGGGATGTGAATGGTAATGATGGCTCAACATGGCTTGCTTCCATTCC
 CTGGGTGAGTGTGCTGCCGTATGAGCTCCAGCTTTGTTCCCTTACTGAGGG
 TTAATTGCGCTTGGGCTAATCATGGTCATAGCTGTTCTGTGAAATTGTTA
 TCCGCTACAATTCCACACATAGAGCGGGAGCATAAGTGAAAGCCTGG
 GTGCCTAATGAGTGAGCTAACATTAATGCGTGCCTACTGCCCTTTC
 CAGTCGGAAACTCTGCTGCGCTTAATGAATCGGCCAACGCCGGGGAG
 AGGCCTTGGCTTGGCCTTCCGCTCTCCGCTACTGACTCGTGCCT
 CGGTGCTTGGCTGCCGAGCGGTATCAGCTCACTCAAAGGCGTAAACGGTTA
 TCCACAGAATCAGGGATAACGCAAGGAAAGAACATGAGCAGAAAAGGGAGCAGAAA
 GGCCAGAACCGTAAAAGGCCGCGTGGCTGGCTTTCCATAGGCTCCGCCCC
 CTGACGAGCATCACAAATCAGCTCAAGTCAGAGGTGGCGAACCCAGAGGA
 CTATAAAGATAACCAGGGTTTCCCCCTGGAGGCTCCCTGCGCTCTGTGTTCC
 GACCCCTGCCCTACGGATACTCTGCCCTTCTCCCTGGGAAGCGTGGCGC
 TTTCTCATAGCTACCGCTGAGGTTCTCAGTGGCTGAGGTCTTCGCTCCAAG
 CTGGCTGAGCTGCAACCCCGTCAAGCAGCAGATTACGCCAGGAAAGGG
 CTATCGTCTGAGTGTGCACTGGCTACAGGACTATCGCCTGCGCAGCCA
 CTGGTAACAGGATTAGCAGCGAGGTATGAGGCTGCTACAGAGTTCTTGAAG
 TGGTGGCTAAACTACGGCTACACTAGAAGGAGCAGTTGGTATCTGCGCTCTGCT
 GAAGCCAGTAACTTGGAAAGAGTTGGTAGCTTGTGATCCGCAAACAAACCA
 CCGCTGAGCTGGGTTTGGTGGCAAGCAGCAGATTACGCCAGGAAAGGG
 GGATCTCAAGAAGATCTTGTATTTCTACGGGGCTGACGCTCAGTGGAAACGA
 AAACCTCACGTTAAGGGATTGGTATGAGATTCTAACAAAGGATCTCACCTAGA
 TCCCTTAAATAAAATGAAGTTTAAATCAATCTAAAGTATATGAGTAAACT
 TGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATCTAGCGATCTGCT
 ATTTCTCATCCATAGTTGCTGACTCCCGTGTAGATAACTACGATACGGG
 AGGGCTTACCATCGGCCAGTGTGCAATGATACCGGAGACCCACGCTCACCG
 GCTCCAGATTATCACAAACAGCCAGGCCAGGGAGCGCAGAAGTGG

US 10,337,001 B2

111

112

- continued

TCCTGCAACTTATCCGCCCATCCAGTCTATTAAATTGTTGCCGGAAAGCTAGAG
 TAAGTAGTTGCCAGTTAATAGTTGCCAACGTTGTGCCATTCTACAGGCATC
 GTGGTGTACGCTCGTGTGGTATGCCATTCTACAGCTCGGGTCCAACGATC
 AAGGCGAGTTACATGATCCCCATGTTGTGCCAAAAAGCGGTTAGCTCCTCGGT
 CTGGTGTGGTGTGGTCAAGAGTAAGTGGCCAGTGTTATCACTCATGGTTATGGCA
 GCACTGACAATACTCTTACTGTCAATGCCCATCGTAAGATGCTTTCTGTGACTGG
 TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCCGGACCGAGTTGCTCT
 GCCCGGGTCAATACGGATAATACCGGCCACATAGC (SEQ ID NO: 20)

EN1mg_ON
target:

GCGTTTCTGGGTGAGAAAAACAGGAAGGCAAATGCCGAAAAAGGGAAATAAGG
 GCGACACGGAAATGTTGAATACTCATACTCTTCTTTCAATTATTGAAGCAT
 TTATCAGGGTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAATA
 AACAAATAGGGGTTCCGGCAGACATTCCCGAAAAGTGCCACCTAAATTGTAAGCG
 TTAATATTGTTAAATTCTCGCTTAAATTGTTAAATCAGCTATTNTTAAAC
 CAAATAGCCGAATCGCAGGAAATCCCTTAAATCAAAGAAATAGACCGGAGATAGG
 GTTGAGTGTGTTCCAGTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA
 ACGTCAAAGGGCAAACCGCTATCAGGGCAGTGCCCACTACGTGAACCATCA
 CCCTAATCAAGTTTGGGTGAGGTGGCTTAAGGCACTAAATCGGAACCTAA
 AGGGAGCCCGATTAGAGCTTGAGGGAAAGCGGGAACGTGGCGAGAAAAGG
 AAGGAAGAAAGCGAAGGGAGCGGGCTAGGGCGCTGGCAAGTGTAGCGGTACAG
 CTGCGCTAACCCACACCCGGCGCTTAATGCCCGCTACAGGGCGGTCCCA
 TTCGCCATTGCGCTGCGGCAACTGTTGGGAAGGGCGATCGGTGCGGGCTTTCG
 TATTACGCCAGCTGGGAAAGGGGATGTGCTGAAGGGATTAAAGTGGTAACG
 CCAGGGTTTCCAGTCAGACGTTGAAAAGCACGGCCAGTGAGCGCGTAATA
 CGACTCACTAAGGGGAATTGGGTACGATCGATGCCCTCTCTACTGCAAC
 CGAAGTCCGGCTCAGGATGTTGCGATGAAAAGTGGTGTGGCGAGCTGG
 GCGCTGGCTGGCGCTGAGTATGAGCTCCAGCTTGTGTTAGGGTGGGG
 TTAATTGCGCTTGGCGTAATCATGGTCATAGCTGTTCTGTGTTAAAGTGT
 TCCGCTCACAACTCCACACACATAGAGCGGAAGCATAAAAGTGTAAAGCCTGG
 GTGCTTAATGAGTGAACCTACATTAATTGCGTTGCCCTCACTGCCCTTTC
 CAGTCGGAAACCTGCGCAGCTGCTTAATGAAATCGCCCAAGCGCGGGAG
 AGGCGGTTTGCCTTGGCGCTTCCGCTTCCGCTACTGACTCGTGC
 CGGTGCTGGCTGGCGCTGCGGCGGTACGCTCACTCAAAGGCAGTAACGGTTA
 TCCACAGAAATCGGGGATAACCGAGGAAAGAAGCATGAGCAGGAAAGGGCAG
 GGCGAGAACCGTAAAGGCGCGTTGCGTGGCTTCCATAGGCTCCGGCCCC
 CTGACGAGCATCACAAATCGACGCTCAAGTCAGAGGTGGCGAACCCGACAGGA
 CTATAAGATAACCGGGTTTCCCCTGGAAAGCTCCCTCGCGCTCTCTGT
 GACCCATGGCTTACCGGATACCTGTCGCTTCTCCCTTGGGAAGCGTGGCG
 TTTCTCATAGCTCACGCTGTAGGTATCTCACTGGTGTAGGTGTTCGTCCAA
 CTGGGCTGTGTCACGAACCCCCCGTCAAGCCGACCGCTGCGCTTATCGGTAA
 CTATCGCTTGTAGTCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCA
 CTGGTAACAGGATTAGCAGAGCGAGGTATGAGGGCTGCTACAGAGTTCTGT
 TGGTGGCTTAACCGCTACACTAGAAGGGACAGTATTGGTATCTGCGCTCTGT
 GAAGCAGGTTACCTCGGAAAGAGTGGTAGCTTGTGATCCGCAAACAAACCA
 CCGCTGTAGCGGTGTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAA
 GGATCTCAAGAAGATCCTTGATCTTCTACGGGCTGACGCTCACTGGAACGCA
 AAACCTCAAGGGATTGGGTTGGTCAAGGATCTACGGGATCTTCACCTAGA
 TCCCTTAAATAAAATGAAGTTAAATCAATCTAAAGTATATGAGTAAACT
 TGGTGTGACAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGT
 ATTCTGCTCATCCATGGCTGACTCCCGTGTGAGATAACTACGATACGGG
 AGGGCTTACCATCGCCCTGCGCAATGATACCGGAGACCCACGTCACCG
 GCTCCAGATTATCAGCAATAACCCAGCCAGCGGAAGGGCGAGCGCAGAGTGG
 TCCGCAACTTATCGCCCTCATCCAGTCTATTAAATTGTTGCCGGAAAGCTAGAG
 TAAGTGGCTGGCAGTTAATAGTTGCCAACGTTGTGCCATTGCTACAGGCA
 GTGGTGTACGCTCGTGTGGTATGCCATTCACTGAGCTCCGGTCCAACGATC
 AAGGCAGGTTACATGATCCCCATGTTGTGCAAAGGCACTTGTGACTGG
 CTCCGATGTGTCAGAGTAAGTGGCGCAGTGTATCACTCATGGTTATGGCA
 GCACGCTACATCTTACTGTCAATGCCCATCCGTAAGATGCTTTCTGTGACTGG
 TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCCGGACCGAGTTGCTCT
 GCCCGGGTCAATACGGATAATACCGGCCACATAGC (SEQ ID NO: 21)

EN1mg_OFF
target:

GCGTTTCTGGGTGAGAAAAACAGGAAGGCAAATGCCGAAAAAGGGAAATAAGG
 GCGACACGGAAATGTTGAATACTCATACTCTTCTTTCAATTATTGAAGCAT
 TTATCAGGGTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAATA
 AACAAATAGGGGTTCCGGCAGACATTCCCGAAAAGGCCACCTAAATTGTAAGCG
 TTAATTTGTTAAATTCTCGCTTAAATTGTTAAATCAGCTATTNTTAAAC
 CAATAGGGCGAAATCGGCAAATCCTTAAATCAAAGAAATAGACCGAGATAGG
 GTTGAGTGTGTTCCAGTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCA
 ACGTCAAAGGGCAAACCGCTATCAGGGCAGTGCCCACTACGTGAACCATCA
 CCCTAATCAAGTTTGGGTGAGGTGGCTTAAGGCACTAAATCGGAACCTAA
 AGGGAGCCCGATTAGAGCTTGAGGGAAAGCGGGAACGTGGCGAGAAAAGG
 AAGGAAGAAAGCGAAAGGGAGCGGGCTAGGGCCTGGCAAGTGTAGCGGTACAG
 CTGCGCTAACCCACACCCGGCGCTTAATGCCGCTACAGGGCGGTCCCA
 TTCGCCATTGCGCTGGCAACTGTTGGGAAGGGCGATCGGTGCGGGCTTTCG
 TATTACGCCAGTGGCGAAAGGGGATGTGCTGAAGGGCAGTGGTAAAGTGGTAACG
 CCAGGGTTTCCGAGTCAGACGTTGAAAACGACGGCCAGTGAGCGCGTAATA
 CGACTCACTAAGGGCGAATTGGGTACGATCGATGCCCTCGTCTTCCGGCG
 CGAACCTGGGGCGCAGGATGTTGCGATGAAGAAGTGGTGTGCGGGTGC
 TGGTGGTGTGCCGGCGCTGATATGCACTGGCTTGTGAGCTTGTGAGGG

US 10,337,001 B2

113

114

- continued

TTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTCCGTGAAATTGTTA
 TCCGCTCACAAATTCCACACAACATAAGCGCGGAACGATAAAGTGTAAAGCCTGGG
 GTGCCTAATGAGTGAGCTAACATCACATAATTGCGTGCCTACTGCCGCTTTC
 CAGTCGGAAACCTGTCGCGCAGTCGCAATTAAATGAATCGCCAACGCCGGGGAG
 AGGCCTTGGCTATTGGCGCTCTTCGCGTCTCGCTACTGACTCGCTGCCT
 CGGTCTGCTCGCGAGCGGTATCAGTCACTCAAAGGCCTAATACGGTTA
 TCCACAGAATCAGGGATAACGCAGAAAGAACATGTGACCAAAGGCCAGCAAA
 GGCCAGGAAACGTAAGGCGCGTGGCTGGCTTTCCATAGGCTCCGCC
 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
 CTATAAAGATAACAGGGCTTCCCCCTGGAAGCTCCCTCGCGCTCTCGTCC
 GACCCCTGCGCTTACCGGATACTGTCGCGCTTCTCCCTCGGGAAAGCGTGGCGC
 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTCGGTAGGTCTCGCTCCAAAG
 CTGGGCTGTGACGGCAACCCCCCTGTCAGCCGACCGCTGCGCTTATCGGTAA
 CTATGCTCTTGAAGTCAACCCGGTAAGACACGACTTATGCCACTGCCAGAGCA
 CTGGTAACAGGATTAGCAGCGAGGTATCTAGGCGTGTACAGAGTTCTGAAG
 TGTTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTGGTATCTGCCTCTGCT
 GAAGGCCAGTTACCTTGGAAAGAAGTGTGAGCTCTGATCCGGCAAAACACCA
 CCGCTGGTAGCGGTGTTTTGGTTGCTAGGCTGAGCTACGGCAGAATTACGGCAGA
 GGATCTCAAGAAGATCTTGTATCTTCTACGGGCTGTACGCTCAGTGGAAACGA
 AAACCTACGTTAAGGGATTGGTATGAGATTATCAAAAGGATCTCACCCTAGA
 TCCTTTAAATAAAATGAAGTTTAATCAATCTAAAGTATATGACTAAACT
 TGGTCTGACAGTTACCAATGCTTAATCAGTCGGGACCTATCTCAGGATCTGTCT
 ATTTCTCATCCATAGTTGCTGACTCCCGTGTAGATAACTACGATACGGG
 AGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGACCCACGCTCACCG
 GCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAAGGGCGAGCCAGAAGTGG
 TCCCTGCAACTTATCCGGCTCCATCCAGTCTTAATATTGTTGCGGGGAAGCTAGAG
 TAAGTAGTTGCGCAGTTAATAGTTGCGCAACGTTGCTGCTTCACTAGGCTACAGGCTAC
 GTGGTGTACGCTCGTGTGGTATGGCTTCATTAGCTCCGGTCTCCAAACGATC
 AAGGCGAGTTACATGATCCCCCATGTTGCTGAAAAAGCGGTTAGCTCTCGGT
 CTGGGATCTGGTGTGAGAGTAAGTGGCCGAGTGTATCACTATGGTTATGGCA
 GCACTGCAATAATTCTTACTGTATGCCATCCGTAAGATGCTTTCTGTGACTGG
 TGAGTACTCAACCAAGTATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTCTT
 GCCCGGTCAATACGGATAATACCGCCACATAGC (SEQ ID NO: 22)

PCDHA4mg_ON
 target : GCGTTTCTGGGTGAGAAAAACAGGAAGGCAAAATGCCGAAAAAGGGATAAGG
 GCGACACGGAAATGTTGAAACTCATACTCTTCTTTCAATATTATTGAGCAT
 TTATCAGGGTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATA
 AACAAATAGGGCTTACCGGCAATTTCCCGAAAAAGTGGCACCTAAATTGTAAGCG
 TTAATTTGTTAAATTCGGGTTAAATTGTTGTTAAATCAGCTCATTTTAAAC
 CAATAGGCCAAATCGGAAATCCCTATAAAATCAAAGAATAGACCGAGATAGG
 GTTGAGTTGTTCCAGTTGGAAACAGACTCCACTTTAAAGAACGTGAACTCCA
 ACGTCAAAGGGAAACCCGCTATCAGGGCATGGGCCACTACGTCGAACCATCA
 CCCTAATCAAGTTGGGGTCGAGGTGGCTAAAGCACTAAATGGGAACCCCTAA
 AGGGAGCCCCGATTAGAGCTTGACGGGAAAGCGCGAACGTGGCGAGAAAGG
 AAGGGAGAAAGCGAACGGAGCGGGCTAGGGCGCTGGCAAGGTAGCGGTACCG
 CTGCGCTAACACCAACCCGCCGCTTAATGCCCGCTACAGGGCGCTCCCA
 TTCGCGGAACTCGGCTGCCACTGTTGGAAAGGGCGATCGTGTGGGGCTTCTCG
 TATTACGGCACTGGGAAAGGGGATGTGCTGCAAGGCGATTAAAGTGGTAACG
 CCAGGGTTTCCAGTCAGCAGCTGTTAAACGACGCCAGTGGCGCGTAATA
 CGAGCTTACAGGGGAATTGGGTACGATCGATGCCCTCGGAACATGGTAAT
 TAAACTTAACGCTTACGATTAGACGAAGGATTGAAATGGGACATTGTTATTCA
 TCTCGAATGATACGGCGTGTATGAGCTCAGCTTGTGAGCTTGTGAGGG
 TTAATTGCGCGTTGGCGTAATCATGGTATAGCTGTTCTGTGAAATTGTTA
 TCCGCTCACAAATCCACACACATACGAGCGGAAGCATAAAGGTAAAGCCTGG
 GTGCGCTAACAGGTTGAGCTAACCTACATGGCTGCGCTACTGCCGCTTTC
 CAGTCGGAAACCTGTCGCGCAGTCGCAATTAAATGCCAACGCCGGGGAG
 AGGCCTGGTGTGCTTGGCGCTCTCCGCTCTCGCTACTGACTCGTGCCT
 CGGTCTGGCTGCGGCTACGCTGCTACGCTAAAGCGGTAAATACGGTTA
 TCCACAGAATCAGGGGAAACGCAAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA
 GGCCAGGAAACCGTAAAGGCCGCGTTGCTGGCTTTCCATAGGCTCCGCC
 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAACCCGACAGGA
 CTATAAAGATAACGGCTTACCGGTTACCGGATACCTGCGCTCTCGTCT
 GACCCCTGCCGCTTACCGGATACCTGTCGCTTCTCCCTCGGGAAAGCGTGGCG
 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTCGGTAGGTCTGCTCC
 CTGGGCTGTGTCAGAACCCCCCTGTCAGCCGACCGCTGCGCTTATCGGTAA
 CTATGCTTCTGAGTCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCA
 CTGGTAACAGGATTAGCAGAGCGAGGTATGAGCTGGCGTCTACAGAGTTCTGAG
 TGGTGGCTTAACCTACGGCTACACTAGAAGGAGCAGTATTGGTATCTGCGCTCTGCT
 GAAGCCAGTTACCTTGGAAAGAAGTGTGAGCTCTGATCCGGCAACAAACCA
 CCGCTGGTAGCGGTGTTTTGGTGTGCAAGCAGCAGATTACGGCAGAAAAAAA
 GGATCTCAAGAAGATCTTGTATCTTCTACGGGCTGTACGCTCAGTGGAAACGA
 AAACCTACGTTAAGGGATTGGTATGAGATTATCAAAAGGATCTCACCTAGA
 TCCTTTAAATTAAAATGAAGTTTAATCAATCTAAAGTATATGAGTAAACT
 TGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGGATCTGTCT
 ATTTCTCATCCATAGTTGCGCTGACTCCCGTGTGAGATAACTACGATACGGG
 AGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGACCCACGCTCACCG
 GCTCCAGATTATCAGCAATAACCCAGCCAGCCGGAAAGGGCGAGCGCAGAAGTGG
 TCTGCGCAACTTATCCGGCTCATCCAGTATTAAATGTTGCGGGAAAGCTAGAG
 TAAGTAGTTGCCAGTTAATGTTGCGCAACGTTGCGCATTGCTACAGGCATC

US 10,337,001 B2

115

116

- continued

GTTGGTGTACGCTCGTCGTTGGTATGGCTTCATTCAAGCTCCGGTCCAAACGATC
AAGGCAGATTACATGATCCCCCATGTTGTGAAAAAAGCGGTTAGCTCCCTCGGT
CTCCGATCGTTGTCAGAAGTAAGTGGCCGAGTGTATCACTCATGGTTATGGCA
GCACTGATAATTCTCTACTGTATGCCATCGTAAGATGCTTTCTGTGACTGG
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCAGCCGAGTTGCTCT
GCCCGGCGTCAATACGGATAATACCGGCCACATAGC (SEQ ID NO: 23)

PCDHA4mg_OFF
target:

GGCGTTCTGGGTGAGAAAAACAGGAAGGCAAATGCCGAAAAAAGGAAATAAGG
GCGACACGGAAATGTGAAACTCATACTCTCCCTTTCAATATTGAAGGCAT
TTATCAGGGTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAATA
AACAAATAGGGGTCGCGCACATTCCCGAAAAGTGCACCTAAATTGTAAGCG
TTAATATTGTTAAATTGCGTTAAATTGTTGTTAAATCAGCTATTGTTAAC
CAATAGGCGAAATCCCTTAAATCAGGAAATAGACCGAGATAGG
GTTGACTGTTGTCAGTTGGAAAGAGTCACATTAAAGAACGTCAGTGGACTCCA
ACGTCAAAGGGCAAAACCGCTATCAGGGCAGTGGCCACTACGTGAACCATCA
CCCTAATCAAGTTGGGTGCGAGGTGCGTAAAGCACTAAATCGGAACCTAA
AGGGGAGCCCGATTAGAGCTTGAGGGAAAGCCGGGAACGTGGCGAGAAAAGG
AAGGGAGAAAGCAGGGAGGAGCGGCGCTGGCAAGTGTAGCGGTACAG
CTGCGCTAACACACACCCGCCGCGCTTAATGCCCGTACAGGGCGCTCCCA
TTCGCCCCATTAGGCTGCCAACTGTTGGAAGGGCAGTCGGTGCGGGCTCTCGC
TATTACGGCAGCTGGGAAAGGGGATGTCTGCAAGGGCATTAAAGTGGGTAACG
CCAGGGTTTCCCAGTACGAGCTGTTAAACAGCGGCAACTGAGCCTGG
CGACTCAACTAGGGGAATTGGGTACGATCGATGCCCTCGGAACGCTGGTGT
TCATCCCAATGCCCTAGATTAGACGAAGGCTGAATGGGGATATTACTCCT
TCTCCAGTGTGCGCGTGTATGAGCTCCAGCTTTGTTCCCTTACTGAGGG
TTAATTGCGCTTGGCTAATCATGGCTAGCTGTTCTGTGAAATTGTTA
TCCGCTCACAACTCCACACACATCGAGCGGAAAGCATAAAGGCTAAAGGCTGG
GTGCCATAATGAGTGAGCTAACATCACATAATTGCGTGCCTACTGCCCGCTTC
CAGTCGGAAACCTGTCGTCGAGCTGCTTAATGAATCGCCAACGCCGGGGAG
AGGGCGTTTGGCTATTGGGCTCTTCCGCTCTCGCTACTGACTCCGCGCT
CGGTCTCGCTCGCGCGAGCGGTACGCTACTCAAAGCGGTAATACGGTTA
TCCACAGAATCAGGGATAACCGAGGAAAGAACATGTGAGCAAAGGGCAGCAAA
GGCCAGAACCGTAAAAGGCCGCTGGCTGGCTTTCCATAGGCTCCGCCCC
CTGACGCTACATGAGCTAACATCACATAATTGCGTGCCTACTGCCCGCTTC
CAGTCGGAAACCTGTCGTCGAGCTGCTTAATGAATCGCCAACGCCGGGGAG
CTATAAGATAACCGGGTTTCCCTGGAAAGCTCCCTGTCGCTCTGGTCT
GACCCTGCCCTACCGATAACCTGTCGCTTCTCCCTCGGAAAGCGTGGCGC
TTTCTCATAGCTCACGCTGAGGTATCTCAAGTCGGTAGGTGCTTCGCTCCAA
CTGGCTGTGTCGACGAAACCCCGTCACTGGCTGACGCTCAGTGGTA
CTATCGCTTGGTCAAGACAGCTTACGGGACTATCGCACTGGCAGCAGCA
CTGGTAACAGGATTAGCAGAGCAGGGTATGAGGCGGTGCTACAGAGTTCTGA
TGGTGGCTTAACCGCTACAGAAGGACAGTATTGGTATCTGCGCTCTGCT
GAAGCCAGTACCTTGGGAAAGAGTTGGTAGCTTGTGATCCGGAAACAAACCA
CCGCTGGTAGCGGTTTTTTGTTGCAAGCAGGAAATACGCCAGAAAAAAA
GGATCTCAAGAAGATCTTGATCTTCTACGGGCTGACGCTCAGTGGAACGA
AAACTCACGTTAAGGGATTGGTATGAGATTATCAAAGGATCTTACCTAGA
TCCCTTTAAATTAAAGTTAAATCAATCTAAAGTATATGAGTAAACT
TGGTCTGCAAGTACCATGCTTAATCAGTGGGACCTATCTCAGCGATCTGTCT
ATTTCGTTCATCCATAGTTGCTGACTCCCGTGTGAGATAACTACGATACGGG
AGGGCTTACCATCTGCCCGCAGTGTGCAATGATAACCGGAGACCCACGCTCACCG
GCTCGAAGATTACGCAATAAACCCAGCGAGCCGGAAAGGGCGAGCCAGTGG
TCCCTGCAACTTACCGCTCCATCCAGCTTAAATTGTCGCGGAAAGCTAGAG
TAAGTAGGTTGCCAGTTAATAGTTGCGAACGTTGTCGCTTACAGGACATC
GTGGTGTACGCTCGCTGGTATGGCTTACCTAGCCTCCGGTCCAAACGATC
AAGGGAGTTACATGCCCTGGTCAAGGTTAGCTCCCTGGTCT
CTCCGATCGTGTGAGAAGTGGCCGAGTGTATCACTCATGGTTATGGCA
GCACTGATAATTCTCTACTGTATGCCATCGTAAGATGCTTTCTGTGACTGG
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCAGCGAGTTGCTCT
GCCCGGCGTCAATACGGATAATACCGGCCACATAGC (SEQ ID NO: 24)

In a 20-uL reaction volume, 50 fmoles of linearized DNA target in the presence of 50 nM sgRNA, 39 nM recombinant purified Cas9 protein (*S. pyogenes*; Agilent) and 10 mM or 0.8 mM MgCl₂ at pH 7.6 was incubated at 37° C. for 30 min. Upon completion, 0.5 uL of RNase It (Agilent) was added, and incubation was continued at 37° C. for 5 min and then at 70° C. for 15 min. Subsequently 0.5 uL of Proteinase K (Mol. Bio. grade, NEB) was added and incubated at 37° C. for 15 min. Aliquots were loaded into a DNA 1000 or DNA 7500 LabChip and were analyzed on a Bioanalyzer 2200, or alternatively were loaded into a Genomic DNA ScreenTape and were analyzed on a TapeStation. The workup steps served to release Cas9 from binding of target DNA, which was assayed for cleavage. Cleavage yields were calculated by the formula: a/(a+b)×100 where a is the sum of the band intensities of the two cleavage products and b is the remain-

ing uncleaved DNA if present. A cleavage percentage of 100% means that all of the target DNA construct was cleaved.

55 A series of guide RNAs were chemically synthesized. The guide RNA oligomers were synthesized on an ABI 394 Synthesizer (Life Technologies, Carlsbad, Calif., USA) using 2'-O-thionocarbamate-protected nucleoside phosphoramidites according to procedures described in Dellingler et al. (2011) *J. Am. Chem. Soc.*, 133, 11540-56. 2'-O-methyl phosphoramidites were incorporated into RNA oligomers under the same conditions as the 2'-O-thionocarbamate protected phosphoramidites. The 2'-O-methyl-3'-O-(diisopropylamino)phosphinoacetic acid-1,1-dimethylcyanoethyl ester-5'-O-dimethoxytrityl nucleosides used for synthesis of thiophosphonoacetate (thioPACE)-modified RNAs were synthesized essentially according to published methods. See

US 10,337,001 B2

117

Dellinger et al. (2003) *J. Am. Chem. Soc.*, 125, 940-50; and Threlfall et al. (2012) *Org. Biomol. Chem.*, 10, 746-54. For phosphorothioate-containing oligomers, the iodine oxidation step after the coupling reaction was replaced by a sulfurization step using a 0.05 M solution of 3-((N,N-dimethylaminomethylidene)amino)-3H-1,2,4-dithiazole-5-thione in a pyridine-acetonitrile (3:2) mixture for 6 min.

All the oligonucleotides were purified using reversed-phase high-performance liquid chromatography (HPLC) and analyzed by liquid chromatography-mass spectrometry (LC-MS) using an Agilent 1290 Infinity series LC system coupled to an Agilent 6520 Q-TOF (time-of-flight) mass spectrometer (Agilent Technologies, Santa Clara, Calif., USA). The yields for the synthesis and purification of the sgRNAs were estimated using deconvolution of mass spectra obtained from LC-MS-derived total ion chromatograms. The chemical synthesis of the 100-mer sgRNAs typically yielded 25-35% full-length product from a nominal 1 micro-mole scale synthesis. Reversed-phase HPLC purification using ion pairing buffer conditions typically gave 20% yield from the crude product with an estimated purity of the final sgRNA in the range of 90% to 95%.

The results are shown in Table 4. "% Target cleaved" indicates the percentage of the target DNA construct which was cleaved. Experiments were run with and without addition of a molar excess of targetless competitor DNA (tcDNA) which potentially competes with the target DNA, so the potential impact of the added nonspecific DNA upon the assay could be seen.

118

TABLE 4-continued

Entry #	[Mg ²⁺] (mM)	tcDNA	% Target cleaved	% CV	% Cleaved vs. CONTROL	% CV CON-TROL
Single-guide scaffold						
26	0.8	N	93%	—	96%	—
27	0.8	Y	94%	5%	100%	5%
28	0.8	N	90%	—	96%	—
29	0.8	Y	0%	5%	0%	5%
30	0.8	N	95%	—	99%	—
31	0.8	Y	94%	5%	100%	5%
32	0.8	N	92%	—	99%	—
33	0.8	Y	84%	5%	99%	5%
34	0.8	N	90%	—	96%	—
35	0.8	Y	94%	5%	100%	5%
36	0.8	N	70%	—	96%	—
37	0.8	Y	0%	5%	0%	5%
38	0.8	N	96%	—	99%	—
39	0.8	Y	59%	5%	100%	5%
Unmodified single-guide RNA (sgRNA)						
40	10	N	93%	—	—	—
41	10	N	94%	—	—	—
42	10	N	94%	—	—	—
43	10	N	92%	—	—	—
44	10	N	90%, 92%	—	—	—
45	10	N	92%	—	—	—
46	10	N	93%	—	—	—
47	0.8	N	86%	—	—	—
48	0.8	N	87%	—	—	—
49	0.8	Y	87%	—	—	—
50	0.8	N	82%	—	—	—
51	0.8	N	92%	—	—	—
52	10	N	60%	—	—	—
53	0.8	N	90%	—	—	—
54	0.8	N	90%	—	—	—
55	0.8	Y	79%	—	—	—
56	0.8	N	79%	—	—	—
57	0.8	N	94%	—	—	—
58	10	N	73%	—	—	—
59	0.8	N	84%	—	—	—
60	0.8	Y	≥85%	—	—	—
61	0.8	Y	89%	—	—	—
62	0.8	N	87%, 82%	—	—	—
63	0.8	N	23%, 22%	—	—	—
2-piece dual-guide scaffold						
Unmodified dual-guide RNA (dgRNA)						
1	0.8	N	99%	—	—	—
2	0.8	Y	99%	5%	—	—
3	0.8	N	96%	—	—	—
4	0.8	Y	100%	5%	—	—
5	0.8	N	96%	—	—	—
6	0.8	Y	0%	5%	—	—
7	0.8	N	99%	—	—	—
8	0.8	Y	100%	5%	—	—
9	10	N	94%, 93%	—	—	—
10	0.8	Y	88%	—	—	—
Fluorophore-coupled dgRNA						
11	10	N	92%, 93%	—	94%, 93%	—
2'OMethyl-modified dgRNA						
40	0.8	Y	87%	—	88%	—
41	0.8	N	88%	—	88%	—
42	0.8	Y	88%	—	88%	—
43	0.8	N	88%	—	88%	—
44	0.8	Y	88%	—	88%	—
45	0.8	Y	88%	—	88%	—
46	0.8	N	88%	—	88%	—
47	0.8	Y	88%	—	88%	—
48	0.8	N	88%	—	88%	—
49	0.8	Y	88%	—	88%	—
50	0.8	Y	88%	—	88%	—
51	0.8	N	88%	—	88%	—
52	0.8	Y	88%	—	88%	—
53	0.8	N	88%	—	88%	—
54	0.8	Y	88%	—	88%	—
55	0.8	Y	88%	—	88%	—
2'OMethyl,3'Phosphorothioate-modified dgRNA						
12	0.8	Y	87%	—	88%	—
13	0.8	Y	87%	—	88%	—
14	0.8	Y	89%	—	88%	—
15	0.8	Y	86%	—	88%	—
2-thioU-modified dgRNA						
16	0.8	N	96%	—	99%	—
17	0.8	Y	95%	5%	99%	5%
18	0.8	N	95%	—	96%	—
19	0.8	Y	100%	5%	100%	5%
20	0.8	N	97%	—	96%	—
21	0.8	Y	0%	5%	0%	5%
22	0.8	N	98%	—	99%	—
23	0.8	Y	99%	5%	100%	5%
24	0.8	N	94%	—	99%	—
25	0.8	Y	83%	5%	99%	5%
DMT-modified sgRNA						
60	0.8	N	78%	—	87%	—
61	0.8	Y	51%	—	87%	—
62	0.8	N	82%	—	87%	—
63	0.8	N	23%, 22%	—	87%	—
64	0.8	N	78%	—	87%	—
65	0.8	Y	76%	—	87%	—
66	0.8	N	65%	—	87%	—
67	0.8	N	81%	—	87%	—
68	0.8	N	85%	—	87%	—
69	0.8	Y	71%	—	87%	—
70	0.8	N	32%	—	87%	—
71	0.8	N	84%	—	87%	—
72	0.8	N	91%	—	87%	—
73	0.8	Y	79%	—	87%	—
74	0.8	N	88%	—	87%	—
75	0.8	N	93%	—	87%	—
76	0.8	N	87%	—	87%	—
77	0.8	Y	79%	—	87%	—
78	0.8	N	89%	—	87%	—
79	0.8	N	88%	—	87%	—
80	0.8	N	3%	—	86%	—
81	0.8	N	5%	—	86%	—
82	0.8	N	89%	—	86%	—
83	0.8	N	68%	—	87%	—
84	0.8	Y	50%	—	87%	—
85	0.8	N	69%	—	87%	—
86	0.8	N	69%	—	87%	—
87	0.8	N	76%	—	87%	—
88	0.8	Y	42%	—	87%	—
89	0.8	N	72%	—	87%	—
90	0.8	N	78%	—	87%	—
91	0.8	N	85%	—	87%	—
92	0.8	Y	51%	—	87%	—
93	0.8	N	82%	—	87%	—
94	0.8	Y	83%	—	87%	—
95	10	N	93%	—	92%	—
96	10	N	93%	—	92%	—

US 10,337,001 B2

119

TABLE 4-continued

Entry #	[Mg ²⁺] (mM)	tcDNA	% Target cleaved	% CV	% Cleaved vs. CONTROL	% CV CON-TROL
Fluorophore-modified sgRNA						
97	10	N	91%, 91%	—	90%, 92%	—
98	0.8	N	86%	—	87%	—
99	0.8	Y	77%	—	87%	—
100	0.8	N	87%	—	87%	—
101	0.8	N	86%	—	87%	—
102	0.8	N	91%	—	87%	—
103	0.8	Y	82%	—	87%	—
104	0.8	N	90%	—	87%	—
105	0.8	N	92%	—	87%	—
106	0.8	N	91%	—	87%	—
107	0.8	Y	82%	—	87%	—
108	0.8	N	90%	—	87%	—
109	0.8	N	91%	—	87%	—
110	0.8	N	92%	—	87%	—
111	0.8	Y	84%	—	87%	—
112	0.8	N	92%	—	87%	—
113	0.8	N	89%	—	87%	—
114	0.8	N	84%, 84%	—	87%, 82%	—
115	0.8	N	12%, 6%	—	23%, 22%	—
116	0.8	N	93%, 90%	—	87%, 82%	—
117	0.8	N	8%, 9%	—	23%, 22%	—
3'Phosphorothioate-modified sgRNA						
118	10	N	95%	—	90%, 92%	—
119	10	N	94%	—	90%, 92%	—
120	10	N	97%	—	90%, 92%	—
121	10	N	94%	—	90%, 92%	—
2'OMethyl-modified sgRNA						
122	10	N	91%	—	94%	—
123	10	N	92%	—	93%	—
124	0.8	N	86%	—	87%	—
125	"	Y	77%	—	87%	—
126	"	N	85%	—	87%	—
127	0.8	N	88%	—	87%	—
128	10	N	92%	—	94%	—
129	0.8	N	83%	—	87%	—
130	0.8	Y	78%	—	87%	—
131	0.8	N	83%	—	87%	—
132	0.8	N	85%	—	87%	—
133	10	N	92%	—	94%	—
134	0.8	N	86%	—	87%	—
135	0.8	Y	78%	—	87%	—
136	0.8	N	83%	—	87%	—
137	0.8	N	88%	—	87%	—
138	10	N	91%	—	94%	—
139	0.8	N	84%	—	87%	—
140	0.8	Y	81%	—	87%	—
141	0.8	N	83%	—	87%	—
142	0.8	N	87%	—	87%	—
143	10	N	89%	—	92%	—
144	0.8	N	91%, 88%	—	87%, 82%	—
145	0.8	N	24%, 25%	—	23%, 22%	—
146	10	N	93%, 92%	—	90%, 92%	—
147	0.8	N	22%	—	87%	—
148	0.8	Y	3%	—	87%	—
149	0.8	N	12%	—	87%	—
150	0.8	N	5%	—	87%	—
151	10	N	0%, 0%	—	90%, 92%	—
152	10	N	0%, 0%	—	90%, 92%	—
153	0.8	N	85%	—	86%	—
154	0.8	N	87%	—	86%	—
155	0.8	N	89%	—	87%	—
156	0.8	Y	78%	—	87%	—
157	0.8	N	84%	—	87%	—
158	0.8	N	93%	—	87%	—
159	0.8	N	90%	—	86%	—
160	0.8	N	90%	—	87%	—
161	0.8	Y	86%	—	87%	—
162	0.8	N	90%	—	87%	—
163	0.8	N	91%	—	87%	—
164	0.8	N	92%	—	90%	—
165	0.8	N	89%	—	87%	—
166	0.8	Y	80%	—	87%	—

120

TABLE 4-continued

Entry #	[Mg ²⁺] (mM)	tcDNA	% Target cleaved	% CV	% Cleaved vs. CONTROL	% CV CON-TROL
5						
167	0.8	N	90%	—	87%	—
168	0.8	N	94%	—	87%	—
169	0.8	N	90%	—	84%	—
170	0.8	Y	≥85%	—	≥85%	—
171	0.8	N	7%	—	84%	—
172	0.8	Y	0%	—	≥85%	—
173	10	N	15%	—	73%	—
174	0.8	N	85%	—	84%	—
175	0.8	Y	75%	—	≥85%	—
176	10	N	86%	—	73%	—
177	0.8	"	0%	—	84%	—
178	0.8	Y	0%	—	≥85%	—
179	10	N	15%	—	73%	—
2'Deoxy-modified sgRNA						
180	10	N	27%, 19%	—	90%, 92%	—
181	10	N	0%, 0%	—	90%, 92%	—
182	10	N	0%, 0%	—	90%, 92%	—
2'Deoxy,3'PACE-modified sgRNA						
183	0.8	N	72%, 77%	—	87%, 82%	—
184	0.8	N	8%, 9%	—	23%, 22%	—
2'OMethyl,3'PACE-modified sgRNA						
25						
185	0.8	N	82%	—	87%	—
186	0.8	Y	72%	—	87%	—
187	10	Y	95%	—	93%	—
188	10	Y	95%	—	94%	—
189	0.8	Y	91%	—	87%	—
190	0.8	Y	84%	—	87%	—
30						
191	0.8	Y	85%	—	87%	—
192	0.8	Y	77%	—	87%	—
193	10	Y	88%	—	94%	—
194	0.8	Y	70%	—	87%	—
195	0.8	Y	56%	—	87%	—
196	0.8	Y	40%	—	87%	—
197	0.8	Y	23%	—	87%	—
198	10	Y	88%	—	93%	—
199	10	Y	89%	—	94%	—
200	0.8	Y	84%	—	87%	—
201	0.8	Y	75%	—	87%	—
202	10	Y	90%	—	93%	—
203	10	Y	90%	—	94%	—
204	0.8	Y	86%	—	87%	—
205	0.8	Y	82%	—	87%	—
206	10	Y	88%	—	93%	—
207	0.8	Y	82%	—	87%	—
208	0.8	Y	78%	—	87%	—
209	10	Y	77%	—	93%	—
210	0.8	Y	71%	—	87%	—
211	"	Y	69%	—	"	—
212	10	N	80%	—	93%	—
213	0.8	N	56%	—	87%	—
214	0.8	Y	41%	—	"	—
215	10	Y	78%	—	93%	—
216	0.8	Y	58%	—	87%	—
217	0.8	Y	44%	—	"	—
218	10	Y	80%	—	93%	—
219	0.8	Y	39%	—	87%	—
220	0.8	Y	13%	—	"	—
221	10	Y	74%	—	93%	—
222	0.8	Y	36%	—	87%	—
223	0.8	Y	19%	—	87%	—
224	10	Y	86%	—	93%	—
225	0.8	Y	84%	—	87%	—
226	0.8	Y	80%	—	"	—
227	10	Y	88%	—	93%	—
228	0.8	Y	83%	—	87%	—
60						
229	0.8	Y	82%	—	87%	—
230	0.8	N	80%	—	87%	—
231	0.8	N	84%	—	87%	—
232	10	N	88%	—	93%	—
233	0.8	N	85%	—	87%	—
234	0.8	Y	73%	—	87%	—
65						
235	10	Y	82%	—	93%	—
236	0.8	Y	89%	—	87%	—

US 10,337,001 B2

121

TABLE 4-continued

Entry #	[Mg ²⁺] (mM)	tcDNA	% Target cleaved	% CV	% Cleaved vs. CONTROL	% CV CONTROL
² O'Methyl,3'Phosphorothioate-modified sgRNA						
237	0.8	Y	76%	—	87%	—
238	10	Y	65%	—	93%	—
239	0.8	Y	84%	—	87%	—
240	0.8	Y	56%	—	87%	—
² O'Methyl,3'PhosphorothioatePACE-modified sgRNA						
241	10	N	92%	—	92%	—
242	0.8	N	84%	—	87%	—
243	0.8	Y	88%	—	87%	—
244	0.8	N	85%	—	87%	—
245	0.8	N	91%	—	87%	—
246	0.8	N	91%	—	84%	—
247	0.8	Y	≥85%	—	≥85%	—
248	0.8	N	84%	—	84%	—
249	0.8	Y	90%	—	89%	—
250	0.8	N	90%, 87%	—	87%, 82%	—
251	0.8	N	16%, 19%	—	23%, 22%	—
252	0.8	N	93%	—	89%	—
253	0.8	N	90%, 90%	—	87%, 82%	—
254	0.8	N	17%, 22%	—	23%, 22%	—
255	0.8	N	93%	—	89%	—
256	0.8	N	91%, 91%	—	87%, 82%	—
257	0.8	N	13%, 16%	—	23%, 22%	—
² O'Methyl,3'PhosphorothioatePACE-modified sgRNA						
258	10	N	89%	—	92%	—
259	0.8	N	84%	—	87%	—
260	0.8	Y	80%	—	87%	—
261	0.8	N	77%	—	87%	—
262	0.8	N	83%	—	87%	—
263	0.8	N	92%	—	87%	—
264	0.8	Y	79%	—	87%	—
265	0.8	N	88%	—	87%	—
266	0.8	N	94%	—	87%	—
267	10	N	74%	—	93%	—
268	0.8	N	11%	—	86%	—
269	0.8	N	15%	—	"	—
270	0.8	N	49%	—	"	—
271	0.8	N	31%	—	"	—
272	0.8	N	91%	—	84%	—
273	0.8	Y	77%	—	≥85%	—
274	0.8	N	90%, 91%	—	87%, 82%	—
275	0.8	N	9%, 8%	—	23%, 22%	—
276	0.8	N	90%	—	84%	—
277	0.8	Y	≥85%	—	≥85%	—
278	0.8	N	86%, 88%	—	87%, 82%	—
279	0.8	N	11%, 7%	—	23%, 22%	—
2-aminoA-modified-sgRNA (including unmodified controls)						
280	0.8	Y	88%, 88%	—		
281	0.8	Y	76%, 75%	—		
282	0.8	Y	87%, 91%	—	88%, 88%	—
283	0.8	Y	90%, 90%	—	76%, 75%	—
284	0.8	Y	85%, 87%	—		
285	0.8	Y	88%, 88%	—		
286	0.8	Y	93%, 96%	—	85%, 87%	—
287	0.8	Y	82%, 79%	—	88%, 88%	—
5-methylIU-modified sgRNA						
288	0.8	N	86%, 83%	—	87%, 82%	—
289	0.8	N	11%, 11%	—	23%, 22%	—
Z base-modified sgRNA						
290	10	N	19%	—	92%	—
291	10	N	93%	—	"	—
sgRNA modified to disfavor misfolding						
292	0.8	N	93%	—	90%	—
293	0.8	N	93%	—	86%	—

The results revealed that guide RNAs containing modifications at specific positions were tolerated by active Cas protein and gRNA:Cas protein complexes, as the modifications did not prevent target-specific cleavage of the on-target polynucleotide. The modifications that were tested and

122

found to be tolerated at specific positions include 2'-O-methylribonucleotide (=2'OMe), 2'-deoxyribonucleotide, racemic phosphorothioate internucleotide linkage, 3'-phosphonoacetate (=PACE), 3'-thiophosphonoacetate (=thioPACE), Z nucleotide, 2-thiouracil, 2-aminoadenine, 5-methyluracil, 5-aminoallyluracil coupled to Cy5 fluorophore, 2-(4-butylamidofluorescein)propane-1,3-diol bis(phosphodiester) linker, and combinations of these.

It is contemplated that the chemical modifications disclosed and tested herein, particularly at the tested positions (as listed in Tables 3 and 4), will be tolerated at equivalent positions in a variety of guide RNAs.

As disclosed herein, chemically modified nucleotides were incorporated into guide RNAs in an effort to improve certain properties. Such properties include improved nucleic acid resistance of the guide RNA (also known as improved stability), reduced off-target effects of a gRNA:Cas protein complex (also known as improved specificity), improved efficacy of gRNA:Cas protein complex when cleaving, nicking or binding a target polynucleotide, improved transfection efficiency, and/or improved organelle localization.

The assay results in Tables 3 and 4 indicate that: (1) In guide RNAs, many positions can tolerate a variety of chemical modifications; (2) 5' and 3' ends of guide RNAs will tolerate a wide variety of end protecting modifications, and such modifications are useful to inhibit exonucleolytic RNA degradation; (3) 2-ThioU can be used to deter off-target interactions involving G-U wobble pairings, thereby increasing the specificity of guide pairing by inhibiting off-target hybridization interactions; (4) 5' Extensions are generally well-tolerated; (5) Surface exposed regions of the guide RNA (as inferred from published crystal structures) are tolerant of extensively modifying U's to 5-methylU's, which potentially make the modified RNA more likely to elude immune responses such as stimulated by unmodified RNA; and (6) For RNA folding, G-C pairs are stronger and more stable than A-U pairs. At least one guide RNA is tolerant of replacing some G-C pairs with 2'-O-methylA-2'-O-methylU pairs that are more stable thermodynamically than unmodified A-U pairs.

More particularly, the present example shows that 2'-O-methyl modifications are tolerated at the 5' and 3' ends of dual-guide RNAs (as shown by entry 12 in Tables 3 and 4) and single-guide RNAs (entries 143-146, 169-170), thus allowing end-protection to stabilize gRNAs against exonucleases. 2'-O-methyl modifications are tolerated at most but not all internal positions, thus allowing stabilization against various nucleases including endonucleases (entries 146, 153-168, 174-179). However, the present example also demonstrates that not every position in guide RNAs will tolerate 2'-O-methyls (as shown by entries 151-152 and 171-173), suggesting that too many consecutive 2'-O-methyl modifications at the 5' end (e.g., 26 or more consecutive 2'-O-methyl-modified nucleotides), or too many 2'-O-methyl modifications of C and U nucleotides downstream (3') of the 5'-terminal 20 mer guide sequence is not well tolerated (e.g., the inhibitory effect of one or both 2'-O-methyluracils at sequence positions +56 and +69 in entries 171-173 as revealed by the positions tested in entries 154-156).

The present example shows that 2'-O-methyl modifications throughout the 20 mer guide sequence are tolerated during in vitro uses in buffer containing 10 mM Mg²⁺ (entry 146), but such extensive modification is not well tolerated under physiological conditions (entries 147-150) as present in cells. Thus, in some embodiments, a gRNA comprising 15 or more, alternatively 17 or more, alternatively 18 or more,

US 10,337,001 B2

123

alternatively 20 2'-O-methyl modifications throughout the 20 mer guide sequence is used for in vitro methods as described herein, such as genomic editing to modify a DNA sequence in vitro, regulating the expression of a gene of interest in vitro, cleaving a DNA target sequence in vitro, and other uses.

The present example shows that extensive incorporation of 2'-deoxy modifications is not well tolerated and can be substantially completely inhibitory (entries 180-182). However, 2'-deoxy modifications can be well-tolerated at some locations (entry 183), therefore such modification can be useful for inhibiting nucleases.

The present example also shows that fluorophore or dye labels are tolerated in every loop of the three known stem-loops in CRISPR-Cas9 guide RNAs (entry 116). Such labels are also tolerated in a 5' overhang on the guide sequence (entry 114), tolerated at additional locations in sgRNAs (entry 114), and tolerated in a loop in tracrRNA used in dual-guide applications (entry 11). In this example, two different types of fluorophore were tested: a phosphodiester-linked fluorophore (no ribose ring) that essentially takes the place of a nucleotide (entries 114 & 116), and a dye label (Cy5) covalently coupled to 5-aminoallylU incorporated in a guide RNA (entry 11).

The present example also shows that Z bases are tolerated in synthetic guide RNAs, particularly as modifications of synthetic guide RNAs in which some C's are replaced with Z bases (entries 290-291). The present example also shows that several other bases are tolerated at various positions, as shown in Tables 3 and 4.

The present example further shows that the 5' and 3' ends of guide RNAs can tolerate a wide variety of end-protecting modifications. Such modifications can be used to inhibit exonucleolytic RNA degradation. Support for the tolerance of such modifications can be found in Hendel et al., *Nat. Biotechnol.* (2015) 55:9, 985-9. Additional support for modifications at the 5' and 3' ends of guide RNAs is provided by entries 143-144, 185-223, 241-257, 258-266, and 272-279 in Tables 3 and 4. In some embodiments, the guide RNA comprises 7 or fewer modified nucleotides at a 5' end or a 3' end or at each of 5' and 3' ends, alternatively 6 or fewer, alternatively 5 or fewer, alternatively 4 or fewer, alternatively 3 or fewer, alternatively 2 or fewer, alternatively 1. dual-guide RNAs can be protected similarly (entries 12-15).

The present example further shows that 2-thioU can be used to deter off-target interactions involving G-U wobble pairings, thereby increasing the specificity of guide sequence pairing by inhibiting off-target hybridization interactions (entries 16-39). One of the base pairs involved in hybridization between the guide RNA and CLTA1 OFF-target 3 (also referred to as "CLTA1 OFF3-target" or "CLTA1 OFF3") is a G-U wobble pair. Replacing the corresponding U in the guide RNA with a 2-thioU reduces cleavage from 100% (entry 8) to 59% (entry 39). Replacing other U's with 2-thioU's (e.g., at sequence position +3 or +9, entries 23 and 31) does not have the same effect, presumably because those U's do not involve G-U wobble pairing when fully hybridized to each of the OFF-target sites tested. Accordingly, 2-thioU can increase target specificity of guide RNAs when off-target sites involve G-U wobble pairing.

The present example also shows that 5'-overhang sequences attached to the guide sequence are generally well-tolerated (see entries 83-95, 114, and 206-223). For example, a bulky dimethoxytrityl (dmt) group at the 5' end was well tolerated (entry 95). The chromatographic properties of dmt can be used to facilitate purification of full-length synthetic RNAs from incompletely elongated byproducts

124

which are generally produced during synthesis. Accordingly, in some embodiments, the synthetic guide RNA comprises a 5'-overhang sequence, for example, comprising a short polynucleotide sequence of 15 or fewer nucleotides which is complementary to the guide sequence and is covalently linked at its 3' end to the 5' end of the guide sequence by a polymeric linker such as a polynucleotide or similar phosphodiester-based linker, in which the linker can be 5 or more consecutive uridine nucleotides, alternatively 6 or 7.

- 10 The present example also shows that surface exposed regions of the guide RNA (as inferred from crystal structures published by others) are tolerant of extensively modifying uracils nucleotides to 5-methyluracils (5-methylU's) (entry 288), which can make the modified RNA more likely to elude immune responses such as stimulated by unmodified RNA. In particular, the 5' and 3' ends of a synthetic guide RNA are potentially immunostimulatory, and the present example shows that 5' and 3' ends are tolerant of 5-methylU modifications (entry 288).
- 15 The present example also shows that a synthetic guide RNA is tolerant of replacing some G-C pairs with 2'-O-methylA-2'-O-methylU pairs which are more stable thermodynamically than unmodified A-U pairs (see the non-terminal-2'-O-methylU and complementary-2'-O-methylA modifications in entries 292-293). This is advantageous because it is known that, for folded RNAs, G-C pairs are stronger and more stable than A-U pairs. Replacement of G-C pairs with such thermostabilized A-U pairs in synthetic guide RNAs allows for improved folding of active structures
- 20 by preventing misfolded structures that involve unintended G-C pair(s), as can be predicted by RNA folding algorithms in common use.

Exemplary Embodiments

35 Exemplary embodiments provided in accordance with the presently disclosed subject matter include, but are not limited to, the claims and the following embodiments:

- 40 A1. A synthetic guide RNA comprising:
a crRNA segment comprising (i) a guide sequence capable of hybridizing to a target sequence, (ii) a stem sequence; and
a tracrRNA segment comprising a nucleotide sequence that is partially or completely complementary to the stem sequence,
wherein the synthetic guide RNA comprises at least one modified nucleotide, and wherein the synthetic guide RNA has gRNA functionality.
- 45 A2. The synthetic guide RNA of embodiment A1, comprising a 2'-deoxy moiety.
- 50 A3. The synthetic guide RNA of embodiment A1 or A2, comprising a 2'-halo moiety selected from 2'-fluoro, 2'-chloro, 2'-bromo and 2'-iodo.
- 55 A4. The synthetic guide RNA of any one of the preceding embodiments, comprising a phosphorothioate group.
- 56 A5. The synthetic guide RNA of any one of the preceding embodiments, comprising a PACE group.
- 60 A6. The synthetic guide RNA of any one of the preceding embodiments, comprising a thioPACE group.
- 65 A7. The synthetic guide RNA of any one of embodiments A2-A6, comprising a 2'-O-methyl moiety.
- A8. The synthetic guide RNA of any one of the preceding embodiments, comprising 2-thiouracil.
- A9. The synthetic guide RNA of any one of the preceding embodiments, comprising a 4-thiouracil.
- 66 A10. The synthetic guide RNA of any one of the preceding embodiments, comprising a 2-aminoadenine.

US 10,337,001 B2

125

A11. The synthetic guide RNA of any one of the preceding embodiments, comprising a hypoxanthine.

A12. The synthetic guide RNA of any one of the preceding embodiments, comprising a 5-methylcytosine.

A13. The synthetic guide RNA of any one of the preceding embodiments, comprising a 5-methyluracil.

A14. The synthetic guide RNA of any one of the preceding embodiments, comprising a 5-aminoallyl-uracil.

A15. The synthetic guide RNA of any one of the preceding embodiments, comprising a Z ribonucleotide.

A16. The synthetic guide RNA of any one of the preceding embodiments, comprising a Z deoxyribonucleotide.

A17. The synthetic guide RNA of any one of the preceding embodiments, comprising a squarate conjugation.

A18. The synthetic guide RNA of any one of the preceding embodiments, comprising a dye linker.

A19. The synthetic guide RNA of embodiment A18, wherein the dye linker is 2-(4-butylamidofluorescein)propane-1,3-diol bis(phosphodiester) linker.

A20. The synthetic guide RNA of any one of the preceding embodiments, comprising a nucleotide with 2'-O-methyl and 3'-phosphorothioate.

A21. The synthetic guide RNA of any one of the preceding embodiments, comprising a nucleotide with 2'-O-methyl and 3'-PACE.

A22. The synthetic guide RNA of any one of the preceding embodiments, comprising a nucleotide with 2'-O-methyl and 3'-thioPACE.

A23. The synthetic guide RNA of any one of the preceding embodiments, comprising a nucleotide with 2'-deoxy and 3'-PACE.

A24. The synthetic guide RNA of any one of the preceding embodiments, comprising a 5-methylcytidine.

A25. The synthetic guide RNA of any one of the preceding embodiments, comprising a methylphosphonate.

A26. The synthetic guide RNA of any one of the preceding embodiments, comprising an ester of PACE, wherein the ester is optionally a methyl ester.

A27. The synthetic guide RNA of any one of the preceding embodiments, comprising a single RNA strand comprising both the crRNA segment and the tracrRNA segment.

A28. The synthetic guide RNA of any one of embodiments A1-A26, comprising two RNA strands, and the crRNA segment and the tracrRNA segment are in different RNA strands.

A29. The synthetic guide RNA of any one of the preceding embodiments, comprising a modified nucleotide at a 5' end, 3' end, or both 5' end and 3' end of each RNA strand.

A30. The synthetic guide RNA of any one of the preceding embodiments, comprising a modified nucleotide in the guide sequence.

A31. The synthetic guide RNA of any one of the preceding embodiments, comprising a modified nucleotide 5' to the guide sequence.

A32. The synthetic guide RNA of any one of the preceding embodiments, comprising a modified nucleotide in the stem sequence.

A33. The synthetic guide RNA of any one of the preceding embodiments, comprising a modified nucleotide in the scaffold region.

A34. The synthetic guide RNA of any one of the preceding embodiments, comprising at least one unnatural, orthogonal base pair in the scaffold region, wherein the base pair is independently selected from isoG-isoC and Z base-P base.

126

A35. The synthetic guide RNA of any one of the preceding embodiments, comprising a 2'-amino group.

A36. The synthetic guide RNA of any one of the preceding embodiments, comprising a phosphorothioate linkage group.

A37. The synthetic guide RNA of any one of the preceding embodiments, comprising a boranophosphonate linkage group.

A38. The synthetic guide RNA of any one of the preceding embodiments, comprising an unlocked nucleic acid modification (ULNA).

A39. The synthetic guide RNA of any one of the preceding embodiments, comprising a locked nucleic acid modification (LNA).

A40. The synthetic guide RNA of any one of the preceding embodiments, comprising an unstructured nucleic acid modification (UNA).

A41. The synthetic guide RNA of any one of the preceding embodiments, comprising a pseudouridine.

A42. The synthetic guide RNA of any one of the preceding embodiments, comprising a 2'-MOE.

A43. The synthetic guide RNA of any one of the preceding embodiments, comprising a 2'-arabino.

A44. The synthetic guide RNA of any one of the preceding embodiments, comprising a 4'-thioribose.

A45. The synthetic guide RNA of any one of the preceding embodiments, comprising a squarate linkage.

A46. The synthetic guide RNA of any one of the preceding embodiments, comprising a triazolo linkage.

A47. A method for cleaving or nicking a target polynucleotide comprising contacting the target polynucleotide with a CRISPR-associated protein and the synthetic guide RNA of any one of the preceding embodiments, and cleaving or nicking the target polynucleotide.

A48. The method of embodiment A47, wherein the cleaving or nicking takes place in vitro.

A49. The method of embodiment A47, wherein the cleaving or nicking takes place in a cell.

A50. The method of embodiment A47, wherein the cleaving or nicking takes place in vivo.

A51. The method of any one of embodiments A47-A50, wherein the CRISPR-associated protein is Cas9.

A52. The method of any one of embodiments A47-A51, wherein the cleaving or nicking results in gene editing.

A53. The method of any one of embodiments A47-A52, wherein the cleaving or nicking results in alteration of gene expression.

A54. A method for binding a target polynucleotide comprising contacting the target polynucleotide with a CRISPR-associated protein and the synthetic guide RNA of any one of the preceding embodiments.

A55. The method of embodiment A54, wherein the CRISPR-associated protein comprises a mutant which does not have a cleavage or nicking activity.

A56. The method of embodiment A54 or A55, wherein the CRISPR-associated protein is a fusion protein comprising a protein component not naturally existing in a CRISPR system.

A57. The method of any one of embodiments A54 to A56, resulting in a change of expression of the target polynucleotide.

A58. The method of any one of embodiments A54 to A57 useful to tag the target polynucleotide.

US 10,337,001 B2

127

Further Exemplary Embodiments

B1. A synthetic guide RNA comprising:

- (a) a crRNA segment comprising (i) a guide sequence capable of hybridizing to a target sequence in a polynucleotide, (ii) a stem sequence; and
- (b) a tracrRNA segment comprising a nucleotide sequence that is partially or completely complementary to the stem sequence,

wherein the synthetic guide RNA comprises one or more modifications, and wherein the synthetic guide RNA has gRNA functionality.

B2. The synthetic guide RNA of embodiment 1, comprising a 2'-O-methyl moiety, a 2'-deoxy moiety, a Z base, a phosphorothioate internucleotide linkage, a phosphonoacetate internucleotide linkage, a thiophosphonoacetate internucleotide linkage, or combinations thereof.

B3. The synthetic guide RNA of embodiment 1 or 2, comprising one or more modifications selected from the group consisting of a 2'-O-methyl nucleotide with a 3'-phosphorothioate group, a 2'-O-methyl nucleotide with a 3'-phosphonocarboxylate group, a 2'-O-methyl nucleotide with a 3'-phosphonoacetate group, a 2'-O-methyl nucleotide with a 3'-thiophosphonocarboxylate group, a 2'-O-methyl nucleotide with a 3'-thiophosphonoacetate group, a 2'-deoxy nucleotide with a 3'-phosphonoacetate group, a 2'-deoxy nucleotide with a 3'-thiophosphonoacetate group, and a Z base.

B4. The synthetic guide RNA of embodiment 1, 2 or 3, comprising one or more modifications selected from the group consisting of a 2'-fluororibosyl, a 2-thiouracil base, a 4-thiouracil base, a 2-aminoadenine base, an hypoxanthine base, a 5-methylcytosine base, a 5-methyluracil base, a methylphosphonate internucleotide linkage, a 5-aminoallyluracil base, a squarate linkage, a triazolo linkage, a dye conjugated to a nucleotide, and combinations thereof.

B5. The synthetic guide RNA of any of the preceding embodiments, comprising a modification selected from the group consisting of a 2'-MOE, 2'-amino, 2'-F-arabino, 2'-LNA, 2'-ULNA, 3'-methylphosphonate, 3'-boranophosphate, 3'-phosphorothioate, 2'-OMe-3'-P(S)₂, 2'-OMe-3'-P(CH₃), 2'-OMe-3'-P(BH₃), 4'-thioribosyl, L-sugar, 2-thiocytosine, 6-thioguanine, 2-aminopurine, pseudouracil, 7-deazaguanine, 7-deaza-8-azaguanine, 7-deazaadenine, 7-deaza-8-azaadenine, 5-hydroxymethylcytosine, 5-hydroxymethyluracil, 5,6-dehydrouracil, 5-propynylcytosine, 5-propynyluracil, 5-ethynylcytosine, 5-ethynyluracil, 5-allyluracil, 5-allylcytosine, 5-allylaminocytosine, P nucleobase, isocytosine (isoC), isoguanine (isoG), UNA, x(A,G,C,T), y(A,G,C,T), abasic nucleotide, PEG, hydrocarbon linker, halo-substituted hydrocarbon linker, heteroatom (O,N,S)-substituted hydrocarbon linker, (keto, carboxy, amido, thionyl, carbamoyl, or thionocarbamoyl)-containing hydrocarbon linker, spermine linker, and combinations thereof.

B6. The synthetic guide RNA of any one of the preceding embodiments, comprising a stability-enhancing modification.

B7. The synthetic guide RNA of any one of the preceding embodiments, comprising at least two modifications; wherein a first modification is a stability-enhancing modification and a second modification is a specificity-altering modification.

B8. The synthetic guide RNA of embodiment 6 or 7, wherein the stability-enhancing modification is located in the guide sequence.

128

B9. The synthetic guide RNA of embodiment 6 or 7, wherein the stability-enhancing modification is located upstream of the guide sequence.

B10. The synthetic guide RNA of embodiment 6 or 7, wherein the stability-enhancing modification is located within the first five and/or the last five nucleotides of the crRNA segment.

B11. The synthetic guide RNA of embodiment 6 or 7, wherein the stability-enhancing modification is located in the tracrRNA segment.

B12. The synthetic guide RNA of embodiment 6 or 7, wherein the stability-enhancing modification is located within the first five and/or the last five nucleotides of the tracrRNA segment.

B13. The synthetic guide RNA of any one of embodiments 6-12, wherein the stability-enhancing modification comprises a 2'-O-methyl moiety, a 2'-fluoro moiety, or a 2'-deoxy moiety.

B14. The synthetic guide RNA of any one of embodiments 6-13, wherein the stability-enhancing modification comprises a phosphorothioate internucleotide linkage, a phosphonoacetate internucleotide linkage, a thiophosphonoacetate internucleotide linkage, a methylphosphonate internucleotide linkage, a boranophosphate internucleotide linkage, or a phosphorothioate internucleotide linkage.

B15. The synthetic guide RNA of any one of embodiments 6-14, wherein the stability-enhancing modification comprises a 3'-phosphonoacetate or a 3'-thiophosphonoacetate.

B16. The synthetic guide RNA any one of embodiments 6-15, wherein the stability-enhancing modification comprises a 2'-O-methyl-3'-phosphorothioate nucleotide, a 2'-O-methyl-3'-phosphonoacetate nucleotide, or a 2'-O-methyl-3'-thiophosphonoacetate nucleotide.

B17. The synthetic guide RNA of any one of embodiments 6-16, wherein the stability-enhancing modification comprises a 2'-fluoro-3'-phosphorothioate nucleotide, a 2'-fluoro-3'-phosphonoacetate nucleotide, or a 2'-fluoro-3'-thiophosphonoacetate nucleotide.

B18. The synthetic guide RNA of any one of the preceding embodiments, comprising a specificity-altering modification.

B19. The synthetic guide RNA of embodiment 18, wherein the specificity-altering modification is located in the guide sequence.

B20. The synthetic guide RNA of any one of embodiment 18 or 19, wherein the specificity-altering modification comprises a 2-thiouracil, a 4-thiouracil or a 2-aminoadenine.

B21. The synthetic guide RNA of any one of embodiments 18-20, wherein the specificity-altering modification comprises a phosphorothioate internucleotide linkage, a phosphonoacetate internucleotide linkage, a thiophosphonoacetate internucleotide linkage, a methylphosphonate internucleotide linkage, a boranophosphate internucleotide linkage, or a phosphorothioate internucleotide linkage.

B22. The synthetic guide RNA of any one of embodiments 18-21, wherein the specificity-altering modification comprises a 3'-phosphonoacetate or a 3-thiophosphonoacetate.

B23. The synthetic guide RNA of any one of the preceding embodiments, comprising a fluorescent dye or a label.

B24. The synthetic guide RNA of any one of the preceding embodiments, comprising one or more isotopic labels.

B25. The synthetic guide RNA of any one of the preceding embodiments, wherein the guide RNA is conjugated to

US 10,337,001 B2

129

an oligonucleotide, an aptamer, an amino acid, a peptide, a protein, a steroid, a lipid, a folic acid, a vitamin, a sugar, or an oligosaccharide.

B26. The synthetic guide RNA of any one of the preceding embodiments, wherein the synthetic guide RNA is a single guide RNA, wherein the crRNA segment and the tracrRNA segment are linked through a loop L.

B27. The synthetic guide RNA of embodiment 26, wherein the loop L comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides.

B28. The synthetic guide RNA of embodiment 26 or 27, wherein the loop L comprises a nucleotide sequence of GNRA, wherein N represents A, C, G, or U and R represents A or G.

B29. The synthetic guide RNA of embodiment 26, 27 or 28, wherein the loop L comprises a nucleotide sequence of GAAA.

B30. The synthetic guide RNA of any one of embodiments 26-29, wherein the loop L comprises one or more modified nucleotides.

B31. The synthetic guide RNA of embodiment 30, wherein the loop L comprises a fluorescent dye.

B32. The synthetic guide RNA of embodiment 31, wherein the dye is conjugated to a 2-(4-butylamido-dye)propane-1,3-diol bis(phosphodiester) linker.

B33. The synthetic guide RNA of any one of the preceding embodiments, wherein the crRNA segment is at the 5' end of the guide RNA.

B34. The synthetic guide RNA of any one of the preceding embodiments, wherein the tracrRNA segment is at the 3' end of the guide RNA.

B35. The synthetic guide RNA of any of the preceding embodiments, wherein the crRNA segment comprises from 25 to 70 nucleotides.

B36. The synthetic guide RNA of any of the preceding embodiments, wherein the guide sequence comprises from 15 to 30 nucleotides.

B37. The synthetic guide RNA of any of the preceding embodiments, wherein the stem sequence comprises from 10 to 50 nucleotides.

B38. The synthetic guide RNA of any of the preceding embodiments, comprising one or more triazolo linkage(s).

B39. The synthetic guide RNA of any of the preceding embodiments, comprising one or more squarate linkage(s).

B40. The synthetic guide RNA of any of the preceding embodiments, wherein the guide RNA comprises a nucleotide composition of:

 M_mN_n

wherein each N independently represents an unmodified nucleotide and each M is selected from a 2'-O-methyl ribonucleotide, a 2'-O-methyl-3'-P(S) ribonucleotide, a 2'-O-methyl-3'-PACE ribonucleotide, a 2'-O-methyl-3'-thioPACE ribonucleotide, and a 2'-deoxynucleotide;

wherein each M is at any position in the sequence of the guide RNA; and

wherein m is an integer between 1 and 220, and n is an integer between 0 and 219, and $50 < m+n \leq 220$.

B41. The synthetic guide RNA of embodiment 38, wherein $m+n < 150$, and each of m and n are independently selected from an integer between 0 and 150, provided that m is not 0.

B42. The synthetic guide RNA of any of the preceding embodiments, wherein the guide RNA comprises a nucleotide sequence of:

 $M_mN_nM'_mN'_nM''_{m''}$
130

wherein each M, M' and M'' independently represent a modified nucleotide and each N and N' independently represent an unmodified ribonucleotide;

wherein any given M is the same or different from any other M, any given N is the same or different from any other N, any given M' is the same or different from any other M', any given N' is the same or different from any other N', any given M'' is the same or different from any other M''; and

wherein m is an integer between 0 and 40, n is an integer between 20 and 130, m' is an integer between 0 and 10, n' is an integer between 0 and 50, m'' is an integer between 0 and 10, provided that $m+m'+m''$ is greater than or equal to 1.

B43. The synthetic guide RNA of any of the preceding embodiments, wherein the crRNA segment comprises a nucleotide sequence of:

 $M_mN_nM'_mN''_{n''}$

wherein M and M' each represent a modified nucleotide and N and N' each represent an unmodified ribonucleotide;

wherein any given M is the same or different from any other M, any given N is the same or different from any other N, any given M' is the same or different from any other M', any given N' is the same or different from any other N'; and

wherein n and n' are each independently selected from an integer between 0 and 50, and wherein m and m' are each independently selected from an integer between 0 and 25, provided that $m+m'$ is greater than or equal to 1.

B44. The synthetic guide RNA of any of the preceding embodiments, wherein the guide sequence comprises a nucleotide sequence of:

 $M_mN_nM'_mN''_{n''}$

wherein M and M' each represent a modified nucleotide and N and N' each represent an unmodified ribonucleotide;

wherein any given M is the same or different from any other M, any given N is the same or different from any other N, any given M' is the same or different from any other M', any given N' is the same or different from any other N'; and

wherein m, n, m', and n' are each independently selected from an integer between 0 and 40, provided that $m+m'$ is greater than or equal to 1.

B45. The synthetic guide RNA of any of the preceding embodiments, wherein the tracrRNA segment comprises a nucleotide sequence of:

 $N_nM_mN'_nM''_{m''}$

wherein M and M' each represent a modified nucleotide and N and N' each represent an unmodified ribonucleotide;

wherein any given M is the same or different from any other M, any given N is the same or different from any other N, any given M' is the same or different from any other M', any given N' is the same or different from any other N'; and

wherein n is an integer between 0 and 130 m is an integer between 0 and 40, and n' is an integer between 0 and 130, and m' is an integer between 0 and 40, provided that $m+m'$ is greater than or equal to 1.

B46. The synthetic guide RNA of any one of embodiments 40-43, wherein m, m', m+m', or $m+m'+m''$ is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

B47. The synthetic guide RNA of any one of embodiments 40-43, wherein m, m', m+m', or $m+m'+m''$ is 1, 2, 3, 4, 5, or 6.

B48. The synthetic guide RNA of any one of embodiments 40-45, wherein n is 16, 17, 18, or 19.

B49. The synthetic guide RNA of any one of embodiments 40-45, wherein n, n', or $n+n'$ is an integer between 75 and 115.

US 10,337,001 B2

131

B50. The synthetic guide RNA of any one of embodiments 40-47, wherein each M is independently selected from the group consisting of a 2'-modified nucleotide, a 3'-modified nucleotide, and combinations thereof.

B51. The synthetic guide RNA of embodiment 48, wherein the 2'-modified nucleotide is selected from the group consisting of a 2'-deoxy nucleotide, a 2'-O-methyl nucleotide, a 2'-fluoro nucleotide, and a 2'-O—C₁₋₃alkyl-O—C₁₋₃alkyl nucleotide.

B52. The synthetic guide RNA of embodiment 48, wherein the 3'-modified nucleotide is selected from the group consisting of a 3'-phosphonoacetate nucleotide and a 3'-thiophosphonoacetate nucleotide.

B53. The synthetic guide RNA of embodiment 48, wherein the combination of the 2'-modified nucleotide and the 3'-modified nucleotide comprises a 2'-O-methyl-3'-phosphorothioate nucleotide, a 2'-O-methyl-3'-phosphonoacetate nucleotide, or a 2'-O-methyl-3'-thiophosphonoacetate nucleotide.

B54. A method for cleaving a target polynucleotide comprising contacting the target polynucleotide with a CRISPR-associated protein and the synthetic guide RNA of any one of the preceding embodiments and cleaving the target polynucleotide.

B55. The method of embodiment 52, further comprising contacting the target polynucleotide with an exogenous CRISPR-associated protein.

B56. The method of embodiment 53, wherein the CRISPR-associated protein is Cas9.

B57. The method of any one of embodiments 52-54, wherein the cleavage results in a functional knockout of a target gene.

B58. The method of any one of embodiments 52-55, further comprising repairing the cleaved target polynucleotide by homology-directed repair with an exogenous or endogenous template polynucleotide.

B59. The method of embodiment 56, wherein the exogenous or endogenous template polynucleotide comprises at least one sequence having substantial sequence identity with a sequence on either side of the cleavage site.

B60. The method of any one of embodiments 52-57, further comprising repairing the cleaved target polynucleotide by non-homologous end joining.

B61. The method of any one of embodiments 56-58, wherein the repairing step produces an insertion, deletion, or substitution of one or more nucleotides of the target polynucleotide.

B62. The method of embodiment 59, wherein the insertion, deletion, or substitution results in one or more amino acid changes in a protein expressed from a gene comprising the target polynucleotide.

B63. The method of any one of embodiments 52-60, wherein the target polynucleotide is contacted with the CRISPR-associated protein and the synthetic guide RNA in vitro.

B64. The method of any one of embodiments 52-61, wherein the target polynucleotide contacted with the CRISPR-associated protein and the synthetic guide RNA is within the genome of a cell in vitro or in vivo.

B65. The method of embodiment 62, wherein the cell is isolated from a multicellular source prior to contacting the target polynucleotide with the CRISPR-associated protein and the synthetic guide RNA.

B66. The method of embodiment 63, wherein the source is a plant, an animal, a multicellular protist, or a fungus.

B67. The method of any one of embodiments 62-64, wherein the cell, or a cell derived therefrom, is returned to

132

the source after contacting the target polynucleotide with the CRISPR-associated protein and the synthetic guide RNA.

B68. A method of modifying a target polynucleotide in a cell comprising introducing into the cell the synthetic guide RNA of any one of embodiments 1-51 and introducing into the cell a CRISPR-associated protein or a nucleic acid that expresses a CRISPR-associated protein in the cell.

B69. The method of embodiment 66, wherein the CRISPR-associated-protein is Cas9.

B70. A method of altering expression of at least one gene product in a cell comprising introducing into the cell the synthetic guide RNA of any one of embodiments 1-51 and further introducing into the cell a CRISPR-associated-protein or a nucleic acid that expresses a CRISPR-associated protein in the cell, wherein the cell contains and expresses a DNA molecule having a target sequence and encoding the gene product.

B71. The method of embodiment 68, wherein the CRISPR-associated-protein is Cas9.

B72. The method of embodiment 69, wherein the CRISPR-associated-protein cleaves the DNA molecule.

B73. A set or library of RNA molecules comprising two or more synthetic guide RNAs of any one of embodiments 1-51.

B74. A kit comprising the synthetic guide RNA of any one of embodiments 1-51 or the set or library of RNA molecules of embodiment 71.

B75. The kit of embodiment 72, further comprising a CRISPR-associated protein or a nucleic acid encoding the CRISPR-associated protein.

B76. The kit of embodiment 73, wherein the CRISPR-associated-protein is Cas9.

B77. The synthetic guide RNA, method or kit of any of the preceding embodiments, wherein the synthetic guide RNA comprises an end modification.

B78. The synthetic guide RNA of any of the preceding embodiments, having a single RNA strand or two separate complementary RNA strands, wherein the guide RNA comprises at least one stability-enhancing modification at both ends of each RNA strand.

C1. A synthetic guide RNA comprising:

(a) a crRNA segment comprising (i) a guide sequence capable of hybridizing to a target sequence in a polynucleotide, (ii) a stem sequence; and

(b) a tracrRNA segment comprising a nucleotide sequence that is partially or completely complementary to the stem sequence,

wherein the synthetic guide RNA comprises one or more modifications, and wherein the synthetic guide RNA has gRNA functionality.

C2. The synthetic guide RNA of embodiment C1, wherein one or more of the modifications comprises a stability-enhancing modification.

C3. The synthetic guide RNA of embodiment C2, wherein one or more of the stability-enhancing modifications is located in the guide sequence.

C4. The synthetic guide RNA of embodiment C2, wherein the stability-enhancing modification comprises a 2'-O-methyl moiety, a Z base, a 2'-deoxy nucleotide, a phosphorothioate internucleotide linkage, a phosphonoacetate (PACE) internucleotide linkage, or a thiophosphonoacetate (thioPACE) internucleotide linkage, or combinations thereof.

C5. The synthetic guide RNA of any of the foregoing embodiments, comprising less than 26 consecutive 2'-O-methyl modified nucleotides at a 5' end of the guide RNA.

US 10,337,001 B2

133

C6. The synthetic guide RNA of any of the foregoing embodiments, comprising a Z base replacing a cytosine in the synthetic guide RNA.

C7. The synthetic guide RNA of any of the foregoing embodiments, comprising at least one 2-thiouracil at a position corresponding to a uridine that can engage in U-G wobble pairing with a potential off-target sequence.

C8. The synthetic guide RNA of any of the foregoing embodiments, comprising one or more modifications selected from the group consisting of a 2'-O-methyl nucleotide with a 3'-phosphorothioate group, a 2'-O-methyl nucleotide with a 3'-phosphonoacetate group, a 2'-O-methyl nucleotide with a 3'-thiophosphonoacetate group, and a 2'-deoxy nucleotide with a 3'-phosphonoacetate group.

C9. The synthetic guide RNA of any of the foregoing embodiments, comprising at least two modifications.

C10. The synthetic guide RNA any of the foregoing embodiments, comprising up to 50 modifications.

C11. The synthetic guide RNA of any of the foregoing embodiments, comprising a single RNA strand or two separate RNA strands, and one or more modifications at a 5' end of each RNA strand, at a 3' end of each RNA strand, or at both a 5' end and a 3' end of each RNA strand.

C12. The synthetic guide RNA of any of the foregoing embodiments, comprising 7 or fewer consecutive modified nucleotides at a 5' end or at a 3' end or at each of 5' and 3' ends.

C13. The synthetic guide RNA of any of the foregoing embodiments, comprising one or more 5-methyluridine nucleotides at one or more of a 5' end, a 3' end or a stem-loop.

C14. The synthetic guide RNA of any of the foregoing embodiments, wherein one or more of the modifications alters base-pairing thermostability.

C15. The synthetic guide RNA of embodiment C14, wherein said one or more modifications enhances the base-pairing thermostability.

C16. The synthetic guide RNA of embodiment C15, wherein said one or more modifications is independently selected from a 2-thiouracil (2-thioU), a 4-thiouracil (4-thioU), a 2-aminoadenine, a 2'-O-methyl, a 2'-fluoro, a 5-methyluridine, a 5-methylcytidine, and a locked nucleic acid modification (LNA).

C17. The synthetic guide RNA of embodiment C15, wherein said one or more modifications decreases the base-pairing thermostability.

C18. The synthetic guide RNA of embodiment C17, wherein said one or more modifications is independently selected from a 2-thiouracil, a 2'-deoxy, a phosphorothioate linkage, a phosphorothioate linkage, a boranophosphonate linkage, a phosphonoacetate linkage, a thiophosphonoacetate linkage, and an unlocked nucleic acid modification (ULNA).

134

C19. The synthetic guide RNA of any of the foregoing embodiments, comprising one or more 2'-O-methylA-2'-O-methylU base pairs.

C20. The synthetic guide RNA of any of the foregoing embodiments, wherein one or more of the modifications is a specificity-altering modification.

C21. The synthetic guide RNA of embodiment C20, wherein the specificity-altering modification is located in the guide sequence.

C22. The synthetic guide RNA of any of the foregoing embodiments, wherein the specificity-altering modification comprises a 2-thiouracil, a 4-thiouracil, a 2-aminoadenine, a 2'-O-methyl, a 2'-fluoro, a LNA, a phosphorothioate linkage, a phosphorothioate linkage, a boranophosphonate linkage, a phosphonoacetate linkage, a thiophosphonoacetate linkage, an ULNA, a 2'-deoxy, a 5-methyluridine, a 5-methylcytidine, or combinations thereof.

C23. The synthetic guide RNA of any of the foregoing embodiments, comprising a fluorescent dye or a label.

C24. The synthetic guide RNA of embodiment C23, wherein the fluorescent dye or a label is bound to a stem-loop of the synthetic guide RNA.

C25. A method for genomic editing to modify a DNA sequence, or regulating the expression of a gene of interest, or cleaving a target polynucleotide comprising: contacting the DNA sequence, the gene of interest, or the target polynucleotide with a CRISPR-associated protein and the synthetic guide RNA of any of the foregoing embodiments, and editing, regulating or cleaving the DNA sequence, the gene of interest or the target polynucleotide.

C26. The method of embodiment C25, wherein the method is performed in vitro, and the synthetic guide RNA comprises 15 or more 2'-O-methyl modifications throughout the guide sequence.

C27. A set or library of RNA molecules comprising two or more synthetic guide RNAs of any of the foregoing embodiments.

C28. A kit comprising the synthetic guide RNA of any of the foregoing embodiments.

C29. An array of RNA molecules comprising two or more, synthetic, guide RNAs of any of the foregoing embodiments.

The foregoing description of exemplary or preferred embodiments should be taken as illustrating, rather than, as limiting, the present invention as defined by the claims. As will be readily appreciated, numerous variations and combinations of the features set forth above can be utilized without departing from the present invention as set forth in the claims. Such variations are not regarded as a departure from the scope of the invention, and all such variations are intended to be included within the scope of the following claims. All references cited herein are incorporated by reference in their entireties.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 130

<210> SEQ ID NO 1
<211> LENGTH: 1368
<212> TYPE: PRT
<213> ORGANISM: *Streptococcus pyogenes*

<400> SEQUENCE: 1

Met	Asp	Lys	Lys	Tyr	Ser	Ile	Gly	Leu	Asp	Ile	Gly	Thr	Asn	Ser	Val
1						5			10				15		

Gly	Trp	Ala	Val	Ile	Thr	Asp	Glu	Tyr	Lys	Val	Pro	Ser	Lys	Lys	Phe
20						25						30			

US 10,337,001 B2

135**136**

-continued

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445

US 10,337,001 B2

137

138

-continued

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

US 10,337,001 B2

139

-continued

140

865	870	875	880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys			
885	890	895	
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp			
900	905	910	
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr			
915	920	925	
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp			
930	935	940	
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser			
945	950	955	960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg			
965	970	975	
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val			
980	985	990	
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe			
995	1000	1005	
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala			
1010	1015	1020	
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe			
1025	1030	1035	
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala			
1040	1045	1050	
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu			
1055	1060	1065	
Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val			
1070	1075	1080	
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr			
1085	1090	1095	
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys			
1100	1105	1110	
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro			
1115	1120	1125	
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val			
1130	1135	1140	
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys			
1145	1150	1155	
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser			
1160	1165	1170	
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys			
1175	1180	1185	
Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu			
1190	1195	1200	
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly			
1205	1210	1215	
Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val			
1220	1225	1230	
Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser			
1235	1240	1245	
Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys			
1250	1255	1260	
His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys			
1265	1270	1275	

US 10,337,001 B2

141**142**

-continued

Arg	Val	Ile	Leu	Ala	Asp	Ala	Asn	Leu	Asp	Lys	Val	Leu	Ser	Ala
1280							1285							1290
Tyr	Asn	Lys	His	Arg	Asp	Lys	Pro	Ile	Arg	Glu	Gln	Ala	Glu	Asn
1295							1300				1305			
Ile	Ile	His	Leu	Phe	Thr	Leu	Thr	Asn	Leu	Gly	Ala	Pro	Ala	Ala
1310						1315						1320		
Phe	Lys	Tyr	Phe	Asp	Thr	Thr	Ile	Asp	Arg	Lys	Arg	Tyr	Thr	Ser
1325						1330					1335			
Thr	Lys	Glu	Val	Leu	Asp	Ala	Thr	Leu	Ile	His	Gln	Ser	Ile	Thr
1340						1345					1350			
Gly	Leu	Tyr	Glu	Thr	Arg	Ile	Asp	Leu	Ser	Gln	Leu	Gly	Gly	Asp
1355						1360					1365			

<210> SEQ ID NO 2

<211> LENGTH: 7

<212> TYPE: PRT

<213> ORGANISM: Simian virus 40

<400> SEQUENCE: 2

Pro	Lys	Lys	Lys	Arg	Lys	Val
1						5

<210> SEQ ID NO 3

<211> LENGTH: 7

<212> TYPE: PRT

<213> ORGANISM: Simian virus 40

<400> SEQUENCE: 3

Pro	Lys	Lys	Lys	Arg	Arg	Val
1						5

<210> SEQ ID NO 4

<211> LENGTH: 16

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 4

Lys	Arg	Pro	Ala	Ala	Thr	Lys	Ala	Gly	Gln	Ala	Lys	Lys	Lys	Lys
1							5				10			15

<210> SEQ ID NO 5

<211> LENGTH: 20

<212> TYPE: PRT

<213> ORGANISM: Human immunodeficiency virus

<220> FEATURE:

<221> NAME/KEY: misc_feature

<223> OTHER INFORMATION: tat protein

<400> SEQUENCE: 5

Gly	Arg	Lys	Lys	Arg	Arg	Gln	Arg	Arg	Pro	Pro	Gln	Pro	Lys	Lys
1													5	10

Lys Arg Lys Val

20

<210> SEQ ID NO 6

<211> LENGTH: 19

<212> TYPE: PRT

<213> ORGANISM: Hepatitis B virus

<220> FEATURE:

<221> NAME/KEY: misc_feature

<223> OTHER INFORMATION: middle S protein, partial

US 10,337,001 B2

143**144**

-continued

<400> SEQUENCE: 6

Pro	Leu	Ser	Ser	Ile	Phe	Ser	Arg	Ile	Gly	Asp	Pro	Pro	Lys	Lys	Lys
1				5				10					15		

Arg Lys Val

<210> SEQ ID NO 7

<211> LENGTH: 24

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 7

Gly	Ala	Leu	Phe	Leu	Gly	Trp	Leu	Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly
1				5				10					15		

Ala Pro Lys Lys Arg Lys Val
20

<210> SEQ ID NO 8

<211> LENGTH: 27

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 8

Gly	Ala	Leu	Phe	Leu	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly
1				5				10					15		

Ala Trp Ser Gln Pro Lys Lys Arg Lys Val
20 25

<210> SEQ ID NO 9

<211> LENGTH: 21

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 9

Lys	Glu	Thr	Trp	Trp	Glu	Thr	Trp	Trp	Thr	Glu	Trp	Ser	Gln	Pro	Lys
1				5				10					15		

Lys Lys Arg Lys Val
20

<210> SEQ ID NO 10

<211> LENGTH: 390

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 10

agaatthaac	tgtggtcaca	tttgcttat	cgactggctt	catctcacag	ctcatcttac	60
gcaagttcga	ttagtatgcc	agtactttc	aatttggttt	aatgtcccg	tgacatgcga	120
gttctgtcga	ccatgtgccg	cggattgaat	tcctcaaggg	tggtgataga	tgctacggtg	180
gtgatgcgca	tgcgctcagt	cctcatctcc	ctcaaggcagg	ccccgctgg	gggtcgagg	240
ccctagtgaa	gccaccaata	tagtggtcg	gtcaagcaac	tgtccacgt	ccaccctcga	300
ggtcgtaca	taaacgtact	aaggcacgag	taaacaagat	cgatagcaag	aacatggtat	360
agactgacgg	agagctcgcc	attagtctga				390

US 10,337,001 B2

145

-continued

146

```

<210> SEQ ID NO 11
<211> LENGTH: 390
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 11
agaatttaac tgtggtcaca tttgctttat cgactggc tt catctcacag ctcatcttac      60
gcaagg tcga tgagtatgcc agtca ct ttc aatttggttt aatgttcccg tgacatgcga    120
gttctgtcga ccatgtgcgcg cggattgaat tcctcaagg gg tgg t gataga tgctacggtg   180
gtgatgcgta tgcactcagt cctcaactcc ctcaaggcagg cgaccctgg ggg t cgg agt    240
cccttagt gaa gccaccaata tagtggcgt gtcaaggcaac tggccacgct ccaccctcga   300
gg t cgt aaca taaacgtact aaggcacgag taaacaagat cgatagcaag aacatggtat 360
agactgacgg agagtcgccc attagtctga                                         390

<210> SEQ ID NO 12
<211> LENGTH: 390
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 12
agaatttaac tgtggtcaca tttgctttat cgactggc tt catctcacag ctcatcttac      60
gcaagg tcga tgagtatgcc agtca ct ttc aatttggttt aatgttcccg tgacatgcga    120
gttctgtcga ccatgtgcgcg cggattgaat tcctcaagg gg tgg t gataga tgctacggtg   180
gtgatgc aat aaattc a g c c t c a t t c c t caaggcagg ggttacttta ggg t cgg agt    240
cccttagt gaa gccaccaata tagtggcgt gtcaaggcaac tggccacgct ccaccctcga   300
gg t cgt aaca taaacgtact aaggcacgag taaacaagat cgatagcaag aacatggtat 360
agactgacgg agagtcgccc attagtctga                                         390

<210> SEQ ID NO 13
<211> LENGTH: 390
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 13
agaatttaac tgtggtcaca tttgctttat cgactggc tt catctcacag ctcatcttac      60
gcaagg tcga tgagtatgcc agtca ct ttc aatttggttt aatgttcccg tgacatgcga    120
gttctgtcga ccatgtgcgcg cggattgaat tcctcaagg gg tgg t gataga tgctacggtg   180
gtgatgc tct ccaggccact cctcatcccc ctcaaggcgg tcccaggcgt ggg t cgg agt    240
cccttagt gaa gccaccaata tagtggcgt gtcaaggcaac tggccacgct ccaccctcga   300
gg t cgt aaca taaacgtact aaggcacgag taaacaagat cgatagcaag aacatggtat 360
agactgacgg agagtcgccc attagtctga                                         390

<210> SEQ ID NO 14
<211> LENGTH: 2838
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:

```

US 10,337,001 B2

147**148**

-continued

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 14

gcgtttctgg	gtgagcaaaa	acaggaaggc	aaaatgccgc	aaaaaaaggga	ataagggcga	60
cacggaaatg	ttgaaatactc	atactcttcc	ttttcaata	ttattgaagc	atttatcagg	120
gttattgtct	catgagcgga	tacatatttgc	aatgtattta	aaaaaaataaa	caaatagggg	180
ttccggegcac	atttccccga	aaagtgcac	ctaaattgtt	agcgtaata	ttttgttaaa	240
attcgcgtta	aattttgtt	aaatcagtc	attttttaac	caataggccg	aaatcgca	300
aatcccttat	aaatcaaaag	aatagaccga	gatagggttg	agtgttgttc	cagtttggaa	360
caagagtcca	ctattaaaga	acgtggactc	caacgtcaaa	gggcgaaaaaa	ccgtctatca	420
gggcgatggc	ccactacgtg	aaccatcacc	ctaatacgat	tttttgggtt	cgagggtggc	480
taaagcacta	aatcggaaacc	ctaaagggag	cccccgattt	agagcttgac	ggggaaagcc	540
ggcgaacgtg	gcgagaaaagg	aagggaaagaa	agcgaaagga	gccccggc	gggcgctggc	600
aagtgttagcg	gtcacgctgc	gctgaaccac	cacacccgc	gctgtttatg	cgccgctaca	660
gggcgcgtcc	cattcgccat	tcaggctgcg	caactgttg	gaagggcgat	cggtgcggc	720
ctcttcgcta	ttacgcccagc	tggcgaaagg	gggatgtgt	gcaaggcgat	taagtgggt	780
aacgccagg	ttttccagt	cacgacgttg	taaaacgacg	gccagtggc	gccccgtata	840
cgactcacta	tagggcgaat	tgggtacgt	cgatgcggcc	tgcaggccca	aagatgtctc	900
ccgcgcgc	tcagtcctca	tctccctaa	gcaggccctg	ctgggtgcact	gaagagccac	960
cctgtgcgcg	tgatatgcag	ctccagcttt	tgttcccttt	agtgggggtt	aattgcgcgc	1020
ttggcgtaat	catggtcata	gctgtttcct	gtgtgaaatt	gttatccgct	cacaattcca	1080
cacaacatac	gagccggaag	cataaaagtgt	aaagcctggg	gtgcctaata	agtggactaa	1140
ctcacattaa	ttgcgttgcg	ctcaactgccc	gctttccagt	cgggaaacct	gtcgtgcag	1200
ctgcattaaat	gaatcggcca	acgcgcgggg	agaggcggtt	tgcgttattgg	gccccgtttcc	1260
gcttcctcgc	tcactgactc	gctgcgcctcg	gtcggtcgcc	tgcggggagc	ggtatcgat	1320
cactcaaagg	cggtataacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	1380
tgagcaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgttttc	1440
cataggctcc	gccccctga	cgagcatcac	aaaaatcgc	gtcaagtc	gagggtggcga	1500
aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtgcgcct	1560
cctgttccga	ccctgcgcgt	tacggatac	ctgtccgcct	ttctcccttc	ggggaaagcgt	1620
gegctttctc	atagetcactg	ctgttaggtat	ctcaagtgcgg	tgttaggtcg	tgcgtccaa	1680
ctggggctgtg	tgcacgaacc	ccccgttca	cccgaccgc	gcccccttac	cggttaactat	1740
cgtcttgagt	ccaacccgg	aagacacgc	ttatgcac	tggcagcgc	cactggtaac	1800
aggattagca	gagcgaggta	tgttaggcgt	gtcacagat	tcttgaagtg	gtggcctaac	1860
tacggctaca	ctagaaggac	agtatttgg	atctgcgc	tgctgaagcc	agttaccc	1920
ggaaaaaaag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cggtggttt	1980
tttggggca	agcagcagat	tacgcgcaga	aaaaaaaggat	ctcaagaaga	tcctttgatc	2040
tttctacgg	ggtctgacgc	tcagtgaaac	aaaaactcac	gttaaggat	tttggtcat	2100
agattatcaa	aaaggatctt	cacctagatc	cttttaattt	aaaaatgaag	ttttaatca	2160
atctaaagta	tatatgagta	aacttggct	gacagttacc	aatgcttaat	cagtggaggca	2220
cctatctcg	cgatctgtct	attcgttca	tccatagttg	cctgactccc	cgtcgtag	2280

US 10,337,001 B2

149**150**

-continued

ataactacga tacgggaggg cttaccatct ggcccagtg ctgcaatgt accgcgagac	2340
ccacgctcac cggctccaga tttatcagca ataaaccaggc cagccggaaag ggccgagcgc	2400
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgtt ccggaaagct	2460
agagtaagta gttcgccagt taatagtttgc cgcaacgttg ttgccattgc tacaggcatc	2520
gtggtgtcac gtcgtcggtt tggtatggct tcattcagct ccggttccca acgatcaagg	2580
c gagttacat gatccccat gttgtc aaa aaagcggtaa gtccttcgg tcctccgatc	2640
gttgtcagaa gtaagttggc cgcaatgttta tcaactatgg ttatggcagc actgcataat	2700
tctcttactg tcatgcattc cgtaagatgc ttttctgtga ctggtagtca ctcaaccaag	2760
tcattctgag aataagtgtat gcccggaccc agttgctttt gcccggcgtc aatacggat	2820
aataccgcgc cacatagc	2838

<210> SEQ ID NO 15
<211> LENGTH: 2838
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 15

gcttttctgg gtgagaaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga	60
cacggaaatg ttgaataactc atactcttcc ttttcaata ttattgaagc atttatcagg	120
gttattgtct catgagcggaa tacatatttgc aatgtatttgc gaaaataaa caaatagggg	180
ttccggegcac atttccccga aaagtgcac ctaaattgttta agcgttaata ttttgttaaa	240
attcgcgttta aattttgtt aaatcagtc attttttaac caataggccg aaatcgccaa	300
aatcccttat aaatcaaaag aatagaccga gatagggttg agtgggttc cagttggaa	360
caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcaaaaaa ccgtctatca	420
ggcgtatggc ccactacgtt aaccatcacc ctaatcaagt tttttgggtt cgagggtccg	480
taaaggacta aatcgaaacc ctaaaggagc cccccgattt agagcttgc ggggaaagcc	540
ggcgaacgttgc gcgagaaagg aagggaaagaa agcgaaaggc gggggcgcta gggcgctggc	600
aagtgttagcg gtcaegctgc gcgtaaccac cacacccgc gcgcttaatg cgccgctaca	660
gggcgegtcc catcgccat tcaggctgcg caactgttgg gaagggcgat cggtgccggc	720
ctcttcgctta ttacgcggc tggcgaaagg gggatgtgtt gcaaggcgat taagtgggt	780
aacgccagggtt ttttccactt caccgttgcg taaaacgcacg gccagtgttgc gcgcttaata	840
cgactcaacta tagggcgaat tgggtacgtt cgtatgcggcc tcgcaggcga aagaggctc	900
ctgttatgcac tcagtcctca actccctcaa gcaggcgacc ctgggtcac tgacaacccg	960
ctcctcgccg tgatatgcag ctccagcttt tggatgggtt aattgcgcgc	1020
ttggcgtaat catggcata gctgtttctt gtgtgaaattt gttatccgtt cacaattcca	1080
cacaacatac gagccggaaag cataaagtgtt aaagcctggg gtgcctaaatg agtggactaa	1140
ctcacattaa ttgcgttgcg ctcactgccc gcttccactt cggaaaccc gtcgtccag	1200
ctgcattaaat gaatcgccca acgcgcgggg agaggcggtt tgcgtattgg gcgctctcc	1260
gcttcctcgcc tcactgactc gctgcgtcg gtcgttgcggc tggcgccggc ggtatcgat	1320
cactcaaagg cggtaatacgtt gttatccaca gaatcaggggg ataacgcagg aaagaacatg	1380
tgagcaaaag gccagcaaaa ggccagggaaac cgtaaaaagg ccgcgttgc ggcgttttc	1440

US 10,337,001 B2

151**152**

-continued

cataggctcc gccccctga cgagcatcac	aaaaatcgac gctaagtca gaggtggcga	1500
aaccgcacg gactataaaag ataccaggcg	tttccccctg gaagctccct cgtgegctct	1560
cctgttccga ccctgcccgt tacggatac	ctgtccgcct ttctcccttc gggaaagcgtg	1620
gcccgttctc atagtcacg ctgttaggtat	ctcagttcg tgtaggtcgt tcgcctcaag	1680
ctgggtgtg tgcacgaacc ccccggtcag	cccgaccgcg ggcgcattatc cggttaactat	1740
cgtctttagt ccaacccggt aagacacgc	ttatcgcac tggcagcagc cactggtaac	1800
aggattagca gagcgaggta tgtaggcggt	gtacagagt tcttgaagt gtggcctaac	1860
tacggctaca ctagaaggac agtattttgt	atctgcgcctc tgctgaagcc agttaccc	1920
ggaaaaagag ttggtagctc ttgatccggc	aaacaaacca ccgctggtag cggtggttt	1980
tttgttgcg acgacgat tacgcgcaga	aaaaaaggat ctcaagaaga tcctttgatc	2040
ttttctacgg ggtctgacgc tcagtggAAC	aaaaactcac gttaaggat tttggtcatg	2100
agattatcaa aaaggatctt cacctagatc	ctttaaatt aaaaatgaag ttttaatca	2160
atctaaagta tatatgagta aacttggct	gacagttacc aatgcttaat cagtggac	2220
cctatctcg cgatctgtct atttcgttca	tccatagttg cctgactccc cgctgttag	2280
ataactacga tacgggaggg cttaccatct	ggcccccagtg ctgcaatgat accgcgagac	2340
ccacgctcac cggctccaga tttatcagca	ataaaccacg cagccggaaag ggccgagcgc	2400
agaagtggtc ctgcaacttt atccgcctcc	atccagtcta ttaattgttg ccggaaagct	2460
agagtaagta gttcggcagt taatagttt	cgcaacgttg ttgccattgc tacaggatc	2520
gtgggttcac gtcgtcggt tggtatggct	tcattcagct ccgggttccca acgtcaagg	2580
cgagttacat gatccccat gttgtcAAA	aaagcggtta gtccttcgg tcctccgatc	2640
gttgtcagaa gtaagtggc cgcagtgtta	tcactcatgg ttatggcagc actgcataat	2700
tctcttactg tcatgccatc cgtaagatgc	tttctgtga ctggtgagta ctcaaccaag	2760
tcattctgag aatagtgtat gccccggaccg	agttgctt gcccggcgtc aatacggat	2820
aataccgcgc cacatagc		2838

<210> SEQ ID NO 16

<211> LENGTH: 2838

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 16

gcgtttctgg gtgagcaaaa acaggaaggc	aaaatgcgc aaaaaaggga ataaggcga	60
cacggaaatg ttgaaatactc atactcttcc	ttttcaata ttattgaagc atttatcagg	120
gttattgtct catgagcggA	tacatatttA aatgtatttA gaaaaataAA caaataggGG	180
ttccgcgcac atttccccga aaagtgcac	ctaaattgtA agcgtaata ttttgtAAA	240
attcgcgtta aattttgtt aaatcagtc	attttttaAC caataggccg aaatggcaA	300
aatcccttat aatcaaaAG aatagaccGA	gatagggttg agtgttGTC cagttggAA	360
caagagtcca ctattaaAGA acgtggactC	caacgtcaAA gggcgaaaaa ccgtctatCA	420
gggcgatggc ccactacgtA aaccatcacc	ctaatacggt tttttgggt cgaggtggcG	480
taaAGcactA aatcgGAacc ctaaaGGGAG	cccccgattt agagcttgc AGGGAAAGCC	540
ggcgaacgtg gcgagAAAGG aagggAAAGG	agcgAAAGGA gggggcgcta gggcgctGGC	600
aagtgtacgc gtcacgctgc	gcttaaccac cacacccGCC ggccttaatG cgccgcTACA	660

US 10,337,001 B2

153**154**

-continued

gggcgcgtcc cattcgccat tcaggctgcg caactgttg gaagggcgat cggtgcggc	720
ctcttcgcta ttacgcgcgc tggcgaaagg gggatgtgtc gcaaggcgat taagtgggt	780
aacgcgcaggg tttcccagt cacgcacgtt taaaacgcacg gccagtgagc gcgcgtataa	840
cgactcacta tagggcgaat tgggtacgt cgatgcggcc tcaggagagg gagccatgct	900
catctccagc ccactcctca tccccctcaa gccggtccca ggctgagagg cttaagcttgc	960
tctttgcgc tgatatgcag ctccagctt tgttccctt agtgagggtt aattgcgcgc	1020
ttggcgtaat catggtcata gctgtttctt gtgtgaaatt gttatccgt cacaattcca	1080
cacaacatac gagccggaaag cataaaagtgt aaaggctggg gtgcctaattg agtgagctaa	1140
ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtgcgtgcag	1200
ctgcattaaat gaatcgccca acgcgcgggg agaggcggtt tgctgttattgg ggcgccttcc	1260
gttccctcgcc tcaactgactc gctgcgcctcg gtcgttccggc tgccggagc ggtatcagct	1320
cactcaaagg cggtaatacgt gttatccaca gaatcaggggg ataacgcagg aaagaacatg	1380
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg cggcggttgc ggcgttttc	1440
cataggotcc gccccctga cgagcatcac aaaaatcgac gtcgaatgtca gaggtggcga	1500
aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgcct	1560
cctgttccga ccctgcccgtt tacggatac ctgtccgcct ttctcccttc gggaaagcgtg	1620
gogcttctc atagtcacg ctgttaggtat ctcaagtccgg tgtaggtcgt tcgctccaa	1680
ctgggctgtg tgcaacgaaacc ccccggttccgc cccgacgcgt ggcgccttac cggtaactat	1740
cgtcttgagt ccaacccgggt aagacacgcac ttatcgccac tggcagcagc cactggtaac	1800
aggattagca gagcggaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaacc	1860
tacggctaca ctagaaggac agtatttttgtt atctgcgcctc tgctgaagcc agttaccc	1920
ggaaaaaagag ttggtagctc ttgatccggc aaacaaaccac ccgctggtag cggtggttt	1980
tttggcttgcgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc	2040
ttttctacgg ggtctgacgc tcagtgaaac gaaaactcactc gttaagggtt tttggtcatg	2100
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaatca	2160
atctaaagta tatatgagta aacttggctc gacagttacc aatgccttaat cagtggggca	2220
cctatctcgat cgtatgtct atttcgttca tccatagttt cctgactccc cgtcgtgttag	2280
ataactacgat tacgggagggtt cttaccatct ggccccagttt ctgcaatgtat accgcggagac	2340
ccacgctcac cggctccaga tttatcagca ataaaccacgc cagccggaaag ggccgagcgc	2400
agaagtggtc ctgcaactttt atccgcctcc atccagttca ttaattgtgtt ccggaaagct	2460
agagtaagta ttgcgcgtt taatagtttgcg cgaacgttgc ttgcattgc tacaggatc	2520
gtgggtgtcact gtcgtcggtt tggtatggct tcattcagctt ccgggtccca acgtcaagg	2580
cgagttacat gatccccat gttgtgcaaa aaagcgggtt gtccttcgg tcctccgtatc	2640
gttgcgttgcgaa gtaagttggc cgcgtgttgc tcactcatgg ttatggcagc actgcataat	2700
tctcttactg tcatgcacatc cgttaagatgc ttttctgtga ctgggtagta ctcaaccaag	2760
tcattctgag aatagtgtat gcggcgaccc agttgtcttgc gcccggcgatc aatacggat	2820
aataccgcgc cacaatgc	2838

<210> SEQ_ID NO 17
<211> LENGTH: 2995
<212> TYPE: DNA

US 10,337,001 B2

155

-continued

156

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<400> SEQUENCE: 17

gcggggcaag agcaactcg	tcgcccata cactattctc	agaatgactt gggtgagta	60
tcaccagtca cagaaaagca	tcttacggat ggcatgacag	taagagaatt atgcagtgt	120
gccataacca tgagtataa	cactgcggcc aacttacttc	tgacaacatg cgaggaccg	180
aaggagctaa ccgcgtttt	gcacaacatg gggatcatg	taactcgct tgatcggt	240
gaaccggagc tgaatgaagc	cataccaaac gacgagcg	acaccacatg gcctgttagca	300
atggcaacaa cgttgcgcaa	actattaact ggcgaactac	ttactctagc ttccggca	360
caattaatag actggatgga	ggcgataaa gttgcaggac	cacttctgcg ctgcggcctt	420
ccggctggct gtttattgc	tgataaatct ggagccgg	tgatcggtatc	480
attgcagcac tggggccaga	tggtaagccc tcccgatcg	tagttatcta cacgacggg	540
atgcaggcaa ctatggatga	acgaaataga cagatcg	agatagggtgc ctcactgatt	600
aagcatttgt aactgtcaga	ccaagtttac tcataatatac	tttagattga tttaaaactt	660
catttttaat taaaaggat	ctaggtgaag atccctttt	ataatctcat gaccaaatac	720
ccttaacgtg agtttcgtt	ccactgagcg tcagaccccg	tagaaaagat caaaggatct	780
tctttagatc cttttttct	gcmcgtatc tgctgcttc	aaacaaaaaa accaccgta	840
ccagcggtgg ttgtttgcc	ggatcaagag ctaccaactc	ttttccgaa ggtactggc	900
ttcagcagag cgcagatacc	aaatactgtt cttctagtgt	agccgtagtt aggccaccac	960
ttcaagaact ctgttagcacc	gccacatacc tcgctctgt	aatcctgtta ccagtggct	1020
cttgccagtg gcgataagtc	gtgtcttacc ggggtggact	caagacgata gttaccggat	1080
aaggcgcagc ggtcgggctg	aacggggggt tcgtgcacac	agcccagctt ggagcgaacg	1140
acctacacccg aactgagata	cctacagcgt gagctatgag	aaagcgcac gcttccgaa	1200
gggagaaagg cggacaggta	tccggtaagc ggcagggtcg	gaacaggaga ggcacgagg	1260
gagctccag gggaaacgc	ctggtatctt tatagtctc	tggttttcg ccacctctga	1320
cttgagcgtc gatTTTGT	atgctcgta gggggccgaa	gcctatggaa aaacgcacgc	1380
aacgcggcct tttacggtt	cctggccctt tgctggcct	ttgctcacat gttttccct	1440
gcgttatccc ctgattctgt	ggataaccgt attaccgc	ttgagtgagc tgataccgt	1500
cgcgcagcc gaacgaccga	gcccgcgcg	tcagtgcg aggaagcggg agagcgcac	1560
atacgaaac cgcctctccc	cgccgcgttgc	ccgattcatt aatgcagctg gcacgacagg	1620
tttcccact ggaaagcggg	cagtgcgc	aacgcaatta atgtgagtt gctcactcat	1680
taggcacccc aggcttaca	ctttatgtt	ccggctcgta tgggtgtgg aattgtgagc	1740
ggataacaat ttcacacagg	aaacagctat	gaccatgatt acgccaagcg cgcaattaac	1800
cctcaactaa gggacaaaa	gctgggtacc	ggggcccccc tcgacaccag ttgcattcg	1860
ttcctgtttg taattgtcca	attctgcag	ccggggggat cggcagatgt agtgtttcca	1920
caggggatcc actagttcta	gagcggccgc	caccgcggtg gagctccaat tcgcccata	1980
gtgagtcgta ttacgcgcgc	tcactggccg	tcgtttaca acgtcgatc tggaaaacc	2040
ctggcggtac ccaacttaat	cgccttcgc	cacatcccc ttgcgcacgc tggcgtaata	2100
gcaagaggc cgcacccgat	cgcccttccc	aacagttgc cagcctgaat ggcgaatgg	2160
aattgtaaac gttatattt	tgttaaaatt	cgcgttaat tttgttaaa tcagtcatt	2220

US 10,337,001 B2

157

-continued

158

ttttaaccaa taggcccggaaa tcggccaaat cccttataaaa tcaaaagaat agaccgagat	2280
agggtttagt gttgttccag tttggAACAA gagtccacta ttaaAGAACG tggactccaa	2340
cgtcaaaaggcg cgaaaaaccg tctatcaggcg cgatggccca ctacgtgaac catcaccccta	2400
atcaagtttt ttgggttcga ggtgcgtaa agcactaaat cggAACCCta aagggagccc	2460
ccgatttaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc	2520
gaaaggagcg ggccgttaggg cgctggcaag tgtagcggtc acgctgcgcg taaccaccac	2580
acccgcgcgc cttaatgcgc cgctacaggcg cgctcagggt ggcactttc gggaaatgt	2640
gcccggaaacc octattttgtt tattttctaa aatacattca aatatgttac cgctcatgag	2700
acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga gtattcaaca	2760
ttccctgtc gcccttattc cctttttgc ggcatttgc cttccgtttttgc ttgctcaccc	2820
agaaacgctg gtgaaagtaa aagatgctga agatcagttt ggtgcacgag tgggttacat	2880
cgaactggat ctcaacagcg gtaagatcct tgagagttt cggccccgaag aacgtttcc	2940
aatgatgagc actttaaag ttctgctatg tggcgccgtt ttatccogta ttgac	2995

<210> SEQ ID NO 18

<211> LENGTH: 2838

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 18

gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga	60
cacggaaatg ttgaatactc atactttcc ttttcaata ttattgaagc atttatcagg	120
gttattgtct catgagcggc tacatatttgc aatgtatttgc gaaaataaaa caaatagggg	180
ttccgcgcac atttcccgaa aaagtgcac ctaaattgtt aacgtttaata ttttgttaaa	240
attcgcgttta aattttgtt aaatcagtc attttttaac caataggccg aaatggcaaa	300
aatcccttat aatcaaaaag aatagacgcg gatagggtt agtgttgc tgcgttggaa	360
caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca	420
gggcgcgttgc ccactacgtg aaccatcacc ctaatcaagt tttttgggtt cgagggtccg	480
taaagcacta aatcggaacc ctaaaggggag cccccgattt agagcttgc ggggaaagcc	540
ggcgttgcgtg gcgagaaaaagg aagggaaaggaa agcgaaaggc gggggcgctt gggcgcttgc	600
aagtgttagcg gtcacgctgc gcttaaccac cacacccgccc gcttaatgc gcccgttaca	660
gggcgcgttcc cattcgccat tcaggctgcg caactgttgc gaagggcgat cggtgcggc	720
ctcttcgtcta ttacgcccgc tggcgaaagg gggatgtgc gcaaggcgat taagtgggt	780
aacggccaggg ttttcccgat cacgcgttg taaaacgcac ggcgttgc ggcgttata	840
cgactcacta tagggcgaat tgggtacgtt cgatgcggcc tcaagagctt cactgagtag	900
gattaagata ttgcagatgt agtgtttcca cagggtggctt cttcagtgc ccaagcgaaac	960
ctgcgtgcgcg tgatatgcagc ctccagctttt tttccctttt agtgcgggtt aattgcgcgc	1020
tggcgtaat catggcata gctgtttccgtt gtgtgaaattt gttatccgtt cacaattcca	1080
cacaacatac gagccggaaag catabaagtgtt aaaggctggg gtcgttataatg agtgcgttac	1140
ctcacatcaa ttgcgttgcg ctcactgccc gctttccagt cggggaaaccc tgcgtgcgcag	1200
ctgcattaaat gaatcgccca acgcgcgggg agaggcggtt tgcgttattgg ggcgttcc	1260

US 10,337,001 B2

159

160

-continued

gcttcctcgc tcactgactc gctgcgcctcg gtcgttcggc tgcggcgagc ggtatcagct	1320
cactcaaagg cggtaatacg gttatccaca gaatcaggggg ataacgcagg aaagaacatg	1380
ttagcaaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgc ggcgttttc	1440
cataggctcc gccccctga cgagcatcac aaaaatcgcac gctcaagtca gaggtggcga	1500
aacccgacag gactataaaag ataccaggcg ttcccccctg gaagctccct cgtgcgtct	1560
cctgttccga ccctgcccgt tacccgatac ctgtccgcct ttctccctc gggaaagcgtg	1620
gcgcgttctc atagctcacg ctgttaggtat ctccagttcgg tgttaggtcgt tcgcctcaag	1680
ctggggctgtg tgcaacgacc ccccggtcag cccgaccgcgt gcgccttatac cgtaactat	1740
cgtcttgagt ccaacccggta aagacacgcac ttatcgccac tggcagcagc cactggtaac	1800
aggatttagca gaggcgggtt tgtagggcggt gctacagagt tcttgaagtg gtggcctaacc	1860
tacggctaca cttagaaggac agtattttgtt atctgcgtct tcgtgaagcc agttacccctc	1920
ggaaaaagag ttggtagctc ttgtatccggc aaacaaacca ccgctggtag cggtggttt	1980
tttgtttgcg acgacgat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgtatc	2040
ttttctacgg ggtctgacgc tcagtgaaac gaaaactcactc gttaagggtt tttggtcatg	2100
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaatca	2160
atctaaagta tatatgagta aacttggctc gacagttacc aatgcctaat cagtggca	2220
cctatctcgat cgatctgtct atttcgttca tccatagtttgcgtactccc cgtcgtag	2280
ataactacgat tacgggaggg cttaccatct ggccccagtg ctgcaatgtat accggcagac	2340
ccacgctcac cggctccaga tttatcagca ataaaccgc cagccggaaag ggccgagcgc	2400
agaagtggtc ctgcaacttt atccgcctcc atccagtcata ttaattttgtt ccggaaagct	2460
agagtaagta gtttgcgcagt taatagtttgcgtt cgcacgttgcgtt tacaggatc	2520
gtgggtgtcac gtcgtcgat ttgtatggct tcattcagtttgcgttccaa acgtcaagg	2580
cgagttacat gatccccat gttgtgcataaa aaagcggtaa gtccttcgg tcctccgtatc	2640
gttgcagaa gtaagttggc cgcagttttgcgtt tcactcatgg ttatggcagc actgcataat	2700
tctcttactg tcatgccttc cgttaagatgc ttttctgtga ctggtagatc ctcaaccaag	2760
tcattctgat aatagtgtat gccccggaccc agttgtctt gcccggcgatc aatacggat	2820
aataccgcgc cacaatgc	2838

<210> SEQ ID NO 19

<211> LENGTH: 2838

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 19

gcgtttctgg gtgagcaaaa acaggaaggc aaaatgcgcg aaaaaaggga ataagggcga	60
cacggaaatg ttgaataactc atactcttcc tttttcaata ttatttgcgtt atttatcagg	120
gttattttgtct catgagcgga tacatatttgcgtt aatgtatggcgtt gaaaaataaa caaataggg	180
ttcccgccac atttcccccga aaagtgcac ctaaatttgcgtt aatgcgttataa tttttgttaaa	240
atccgcgttataa aatttttgtt aatgcgttataa attttttgcgtt aatgcgttataa tttttgttaaa	300
aatcccttat aatcaaaaag aatagaccga gatagggttgcgtt aatgcgttataa tttttgttaaa	360
caagagtcca ctatcaaaaag aatagaccga gatagggttgcgtt aatgcgttataa tttttgttaaa	420
gggcgttgcgtt aatgcgttataa tttttgttaaa	480

US 10,337,001 B2

161**162**

-continued

taaaggacta aatcggaacc ctaaaggagccccgattt agagcttgcacggaaagcc	540
ggcgaacgtg gcgagaaaagg aaggaaagaa agcgaaaggagccggggcgcttggc	600
aagttagcg gtcacgctgc gcgtaaccac cacacccgccgcttaatcgccgctaca	660
ggcgcggtcc cattcgccat tcaggctcgcaactgttgggaaggcgatcggtcgccc	720
cttcgccta ttacgcccage tggcgaaagg gggatgtgct gcaaggcgat taagttgggt	780
aacgcgcagggtttcccgat cacgacgttg taaaacgacg gccagtgagc gcgctata	840
cgaactacta tagggcgaaat tgggtacgat cgatgcggcc tctctgaata gagttggaa	900
gagatgcata caacatatgt agtatttcca cagggaatac aatggacaaa tgacctcaag	960
agcaggcgcg tgatatgcag ctccagctt tgttccctt agtgagggtt aattgcgcgc	1020
tggcgtaat catggtcata gctgtttctt gtgtgaaatt gttatccgct cacaatttcca	1080
cacaacatac gagccggaaag cataaagtgt aaagcctggg gtgcctaatcgatgagctaa	1140
ctcacattaa ttgcgttgcg ctcactgccc gcttcccgat cgggaaacct gtcgtgcag	1200
ctgcattaaat gaatcgccaa acgcgcgggg agaggcggtt tgcttattgg ggcgtttcc	1260
gtttccctcg tcactgactc gctgcgcgtcg gtcgttccggc tgccggcgagc ggtatcgat	1320
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg	1380
tgagcaaaag gccagaaaa ggccaggaac cgtaaaaagg ccgcgttgc ggcgttttc	1440
cataggctcc gccccctga cgagcatcac aaaaatcgac gtcagaatca gaggtggcga	1500
aacccgcac gactataaag ataccaggcg tttccctgtt gaagctccct cgtgegctct	1560
cctgttccga ccctgcccgtt taccggatac ctgtccgcct ttctcccttc gggagcggt	1620
ggcgcttctc atagctcagc ctgttaggtat ctctggatcg tggatgtcg tgcgtccaa	1680
ctgggctgtg tgcaegaacc ccccggttccag cccgacccgt ggccttatac cgtaactat	1740
cgtctttagt ccaacccggta aagacacgc ttatcgccac tggcagcagc cactggtaac	1800
aggattagca gagcgaggta tggatggcggt gctacagatg tcttgaagtg gtggctaac	1860
tacggctaca ctagaaggac agtatttggt atctgcgcctc tgctgaagcc agttaccc	1920
ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggttttt	1980
tttgttgcg agcagcagat tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc	2040
ttttctacgg ggtctgacgc tcagtgaaac gaaaactcac gtttggat tttggatcg	2100
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaatca	2160
atctaaagta tatatgagta aacttggctc gacagttacc aatgtttaat cagtggcgca	2220
cctatctcg cgatctgtct atttcgttca tccatagttt cctgactccc cgctgttag	2280
ataactacgta tacgggagggtt cttaccatct ggccccagtg ctgcaatgtat accgcgagac	2340
ccacgctcac cggctccaga tttatcagca ataaaccaggc cagccggaaag ggccgagcgc	2400
agaagtggtc ctgcaactttt atccgcctcc atccagtctttaattgttgc cggggagct	2460
agagtaagta gttcgccagt taatagtttgc gcaacgttg ttgcccattgc tacaggcatc	2520
gtgggtgtcactcgtcgat tggatgggtt tcattcagct cgggttccca acgtcaagg	2580
cgagttacat gatccccat gttgtgcaaa aaagcggtt gtccttcgg tcctccgtat	2640
gttgcgaa gtaagtggc cgcagtgatc tcaactcatgg ttatggcagc actgcataat	2700
tctcttactg tcatgcccattc cgtaagatgc tttctgtga ctggtgagatc ctcaaccaag	2760
tcattctgatc aatagtgtat gccccggaccc agttgttccat gcccggcgatc aatacggat	2820

US 10,337,001 B2

163

164

-continued

aataccgcgc cacatagc	2838
<210> SEQ ID NO 20	
<211> LENGTH: 2838	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: synthetic construct	
<400> SEQUENCE: 20	
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga	60
cacggaaatg ttgaatactc atacttcc ttttcaata ttattgaagc atttatcagg	120
gttattgtct catgagcgga tacatattt aatgtattt gaaaaataaa caaatagggg	180
ttccgcgcac atttccccga aaagtgcac cttaatttga agcgttaata ttttgttaaa	240
attcgcgtta aattttgtt aaatcagctc attttttaac caataggccg aaatcgcaaa	300
aatcccttat aatcaaaaag aatagaccga gatagggttg agtgttgc cagttggaa	360
caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca	420
gggcgatggc ccactacgtg aaccatcacc ctaatcaagt ttttggggt cgaggtgccc	480
taaagcacta aatcggaacc cttaagggag cccccgattt agagcttgc ggggaaagcc	540
ggcgaacgtg gcgagaaaagg aagggaaagaa agcggaaagga gcggggcgcta gggcgctggc	600
aagtgttagcg gtcacgctgc gctgttgc accatccac cacaatcgcc gctttttatgc cggcgctaca	660
gggcgcgtcc cattgcctat tcaggctgcg caactgttg gaagggcgat cgggtgcggc	720
ctcttcgcta ttacgcccgc tggcgaaagg gggatgtgct gcaaggcgat taagttgggt	780
aacgcccaggg tttcccaagt cacgacgttg taaaacgacg gccagtgcgc ggcgttaata	840
cgactcacta tagggcgaat tgggtacgt cgatgcggcc tcggggcagat gcaggataaa	900
gagggatgtg aatggtaatg atggcttcaa catggcgctt gcttttattt ccctgggtgt	960
agtctgcgcg tgatatgcag ctccagcttt tttccctttt agtgggggtt aattgcgcgc	1020
ttggcgtaat catggtcata gctgtttctt gtgtgaaatt gttatccgtt cacaatttcca	1080
cacaacatac gagccggaaag cataaaagtgt aaagcctggg gtgcctaatg agtggactaa	1140
ctcacattaa ttgcgttgcg ctcaactgccc gctttccagt cggggaaacct gtcgtgcag	1200
ctgcattaaat gaatcgccca acgcgcgggg agaggcggtt tgctgttgg ggcgttcc	1260
gcttcctcgc tcactgactc gctgcgtcg gtcgttgcg tgccggcagc ggtatcgat	1320
cactcaaagg cggtaatacgt gttatccaca gaatcaggggg ataacgcagg aaagaacatg	1380
tgagcaaaaag gccagcaaaa ggccagggaaac cgtaaaaagg cccgcgttgc ggcgttttc	1440
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga	1500
aacccgacag gactataaag ataccaggcg tttcccttg gaagctccct cgtgegetct	1560
cctgttccga ccctgcgcgtt taccggatac ctgtccgcct ttctcccttc gggagcggt	1620
ggcgtttctc atagctcacg ctgttaggtat ctcaigtccg tgtaggtcg tgcgtccaa	1680
ctgggctgtg tgacacgacc cccgcgttgc cccgaccgc ggcgccttac cggtaactat	1740
cgtcttgagt ccaacccgggt aagacacgcac ttatcgccac tggcagcagc cactggtaac	1800
aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagt gtcgttgc	1860
tacggctaca otagaaggac agtattttgtt atctgcgttgc tgctgaagcc agttaccc	1920
ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggttttt	1980
tttggggca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc	2040

US 10,337,001 B2

165

-continued

166

ttttctacgg ggtctgacgc tcagtggAAC	2100
aaaaactcac gttaaggat tttggtcATG	
agattatcaa aaaggatctt cacctAGATC	2160
cTTTAAATT AAAAATGAAG TTTAAATCA	
atctaaAGTA tatatGAGTA aactGGTCT	2220
gacAGTTACc aATGCTTAAT cAGTGAGGC	
cCTATCTCAG CGATCTGTCT ATTCGTTCA	2280
TCCATAGTTG CCTGACTCCC CGTCGTGTA	
ataactacga tacGGGAGGG CTTACCACtC	2340
GGCCCCAGTG CTGCAATGAT ACCGCGAGAC	
CCACGTCAC CGGCTCCAGA TTTATCAGCA	2400
ATAAAACCAGC CAGCCGGAAG GGCCGAGCGC	
agaAGTGGTC CTGCAACTT ATCCGCTCC	2460
ATCCAGTCTA TTAATTGTTG CCGGGAAAGCT	
AGAGTAAGTA GTTCGCCAGT TAATAGTTG	2520
CgcaACGTTG TTGCCATTGc TACAGGCATC	
GtGGTGTcAc GCTCGTCGTT TGGTATGGCT	2580
TcATTcAGCT CCGGTTCCCA ACgATCAAGG	
CGAGGTTACAT GATCCCCCAT GTTGTGCAA	2640
AAAGCggTTA GTCCTTcGG TCTCCGATC	
GTTGTCAGAA GtaAGTGGC CGCAGTGTa	2700
TCACTCATGG TTATGGCAGC ACTGCATAAT	
TCTCTTACTG TCACTGCATC CGTAAGATGC	2760
TtttCTGTGA CTGGTGTAGA CTCAACCAAG	
TcATTCTGAG AATAGTGTAT GCGGGGACCG	2820
AGTTGCTTCTT GccccGGGAGTC AATAcGGGAT	
AATAACCGCGC CACATAGC	2838

<210> SEQ ID NO 21

<211> LENGTH: 2838

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 21

gcgtttctgg gtgagcaaaa acagGAAGGC AAAATGCCGC	60
aaaaaAGGGA ATAAGGGCGA	
cacggaaatg ttGAATAACTC atactCTTCC	120
ttttcaata ttATTGAAGC ATTtATCAGG	
gttattgtct catgAGCGGA tacatATTG	180
aatgtATTtA gAAAATAAA CAAATAGGGG	
ttccGCGCAC atttCCCGA aaAGTGCAC	240
ctaaATTGTA AGCgttaata ttttGTTAAA	
attcgeGTTA aTTTTGTT aaATCAGTC	300
atTTTTAAc CAATAGGCCG AAATCGGCAA	
aATCCCTTAT aaATCAAAG AATAGACCGA	360
gatAGGGTTG AGTGTGTTc CAGTTGAA	
caAGAGTCCA CTATTAAGA ACgtggACTC	420
caACgtCAA GGGCGAAAAA CGTCTATCA	
gggcgatGGC CCACTACGT	480
aaccATCACC CTAATCAAGT	
ttttGGGGT CGAGGTGCG	
tttttGGGGT CGAGGTGCG	540
taaAGCActA aATCGGAACC	
CTAAAGGGAG CCCCGATTt AGAGCTGAC	600
GGGGAAAGCC	
ggcgaACGTG GCGAGAAAGG	
aAGGGAAGAA AGCGAAAGGA	660
gCGGGCGCTA GGGCGCTGGC	
aAGTGTAGCG GTCACGCTGC	
gCGTAACCAC CACACCCGCC	720
CGCTTAATG CGCCGCTACA	
gggcgcgtcc CATTGCCCCAT	
TcAGGCTCG CAACTGTGG	780
GAAGGGCGAT CGGTGCGGGC	
CTCTTCGCTA TTACGCCAGC	
TGGCGAAAGG GGGATGTGCT	840
GCAAGGGCGAT TAAGTTGGGT	
AACGCCAGGG TTTTCCAGT	
CACGACGTT TAAAACGACG	900
GCCAGTGAGC GCGCGTAATA	
CGACTCActA TAGGGCGAAT	960
TGGGTACGAT CGATGCGGCC	
TCTCCTTAC TGCAGCCAA	
GTCCGGCCTC AGGATGTTGT	1020
CGATGAAAAA GTTGGTGGT	
CGGTGCAAGT GGGCCGCTGG	
TGCGGGCGCG TGATATGCAg	1080
CTCCAGCTT TGTCCCTT	
AGTGAAGGTT ATTGCGCGC	
TGCGGTGATA CATGGTCATA	
GCTGTTCCT GTGTGAAATT	1140
GTTATCCGCT CACAATTCCA	
CACAACATAC GAGCCGGAAG	
CATAAAAGTGT AAAGCCTGGG	1200
GTGCCTAATG AGTGAAGCTAA	
CTCACATTAA TTGCGTTGCG	
CTCACTGCCC GCTTCCAGT	
CGGGAAACCT GTCGTGCCAG	

US 10,337,001 B2

167**168**

-continued

ctgcattaat	gaatcgccca	acgcegcgggg	agaggcgggtt	tgcgtattgg	gcgctctcc	1260
gcttcctcgc	tcactgactc	gctgcgcctcg	gtcggttcggc	tgcggegagc	ggtatcagct	1320
cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	1380
tgagcaaaag	gccagaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgttttc	1440
cataggctcc	ccccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gagggtggcga	1500
aacccgcac	gactataaaag	ataccagcgg	tttccccctg	gaagctccct	cgtgegcct	1560
cctgttccga	ccctgccgt	tacggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	1620
gwgcttctc	atagctcacg	ctgttaggtat	ctcagttcgg	tgttagtcgt	tcgctccaag	1680
ctgggctgtg	tgcacgaacc	ccccgttcag	cccgaccgct	gccccttatac	cgttaactat	1740
cgtctttagt	ccaacccgg	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	1800
aggattagca	gagcggaggt	tgtaggcgg	gctacagagt	tcttgaagt	gtggcctaac	1860
tacggctaca	ctagaaggac	agtatgggt	atctgcgctc	tgctgaagcc	agttaccc	1920
ggaaaaaagag	ttggtagctc	ttgatccggc	aaacaaaacc	ccgctggtag	cggtggttt	1980
tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	2040
tttctacgg	ggtctgacgc	tcagtgaaac	aaaaactcac	gttaaggat	tttggtcatg	2100
agattatcaa	aaaggatctt	cacctagatc	ctttaaatt	aaaaatgaag	ttttaatca	2160
atctaaagta	tatatgagta	aacttggtct	gacagttacc	aatgcttaat	cagtgaggca	2220
cctatctcg	cgatctgtct	attcgttca	tccatagtt	cctgactccc	cgtcgttag	2280
ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	2340
ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	2400
agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgtg	ccggaaagct	2460
agagtaagta	gttcgcccgt	taatagttt	cgcaacgttg	ttgccattgc	tacaggcatc	2520
gtgggtgtcac	gtcgtcg	ttgtatggct	tcattcagct	ccggttccca	acgatcaagg	2580
cgagttacat	gatccccat	gttgcgaaa	aaagcggta	gtccttcgg	tcctccgatc	2640
gttgcgaa	gtaagtggc	cgcgtgtt	tcactcatgg	ttatggcagc	actgcataat	2700
tctcttactg	tcatgccatc	cgttaagatgc	tttctgtga	ctggtgagta	ctcaaccaag	2760
tcattctgag	aatagtgtat	gcggcgaccc	agttgcctt	gcccggcgtc	aatacggat	2820
aataccgcgc	cacatagc					2838

<210> SEQ ID NO 22

<211> LENGTH: 2838

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 22

gcgttctgg	gtgagaaaa	acaggaaggc	aaaatgcgc	aaaaaaggga	ataagggcga	60
cacggaaat	ttgaataactc	atactctcc	ttttcaata	ttattgaagc	atttatcagg	120
gttattgtct	catgagcgg	tacataattt	aatgtattt	aaaaaataaa	caaatagggg	180
ttccgcgcac	atttccccga	aaagtgcac	ctaaattgt	agcgttaata	ttttgttaaa	240
atccgcgtt	aattttgc	aaatcagctc	atttttaaac	caataggccg	aaatcgccaa	300
aatccctat	aaatcaaaag	aatagaccg	gatagggtt	agtgtgtc	cagtttgaa	360
caagagtcca	ctattaaaga	acgtggactc	caacgtcaaa	ggcgaaaaaa	ccgtctatca	420

US 10,337,001 B2

169

-continued

170

gggcgatggc ccactacgtg aaccatcacc ctaatcaagt ttttggggt cgaggtgccg	480
taaagacta aatcgaaacc ctaaaggag cccccgatt agagctgac ggggaaagcc	540
ggcgaacgtg gcgagaaagg aagggaaagaa agcgaaagga gggggcgcta gggcgctggc	600
aagtgttagcg gtcacgctgc gcttaaccac cacacccgccc gcgcttaatg cgccgctaca	660
gggcgctcc cattcgccat tcaggctgcg caactgttgg gaagggcgat cggtgcggc	720
ctcttcgcta ttacgcgcg tggcgaaagg gggatgtgct gcaaggcgat taagtgggt	780
aacgcacaggg tttcccaagt caacgacgtt gaaaacgacg gccagtgagc gcgcgtataa	840
cgactcaacta tagggcgaat tgggtacgat cgatgcggcc tcgtcccttc gccggccgaa	900
ctcggccgc aggatgttgt cgatgaagaa gttggtgatg cggtgcgggt gctggtggtt	960
gcggggcgcg tgatatgcag ctccagctt tggcccttt agtgagggtt aattgcgcgc	1020
ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca	1080
cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg agtgagctaa	1140
ctcacattaa ttgcgttgcg ctcaactgccc gcttccagtg cgggaaacct gtcgtgcag	1200
ctgcattaaat gaatcgccca acggcgcccc agaggcggtt tgcgtattgg ggcgtctcc	1260
gcttcctcgc tcaactgactc gctgcgcctcg gtcgttcggc tgcggcgagc ggtatcagct	1320
cactcaaagg cggtaatacg gttatccaca gaatcaggggg ataacgcagg aaagaacatg	1380
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg cggcgttgc ggcgttttc	1440
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga	1500
aacccgacag gactataaag ataccaggcg tttccctcg gaagctccct cgtgegctct	1560
cctgttccga ccctgcgcct tacggatac ctgtccgcct ttctcccttc gggaaagcgtg	1620
ggttttctc atagtcacg ctgttaggtat ctcaagttcg tgcgttcgt tgcgtccaa	1680
ctgggctgtg tgcaegaacc ccccggtcag cccgacgcgt ggccttatac cggtaactat	1740
cgtcttgagt ccaacccggta aagacacgac ttatcgccac tggcagcagc cactggtaac	1800
aggattagca gagcgaggta tgcgttcgtt gctacagagt tcttgaagtg tgcgtccaa	1860
tacggctaca ctagaaggac agtattttgtt atctgcgcctc tgcgtaaaggcc agttaccc	1920
ggaaaaaagag ttggtagctc ttgatccggc aaacaaacca cggctggtag cgggtggttt	1980
tttgggtgc aacggatctt tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc	2040
tttctacgg ggtctgacgc tcaactggaaac gaaaactcac gttaagggtt tttgggtatg	2100
agattatcaa aaaggatctt cacctagatc cttttaatt aaaaatgaag ttttaatca	2160
atctaaagta tatatgagta aactggtct gacagttacc aatgcctaat cagtggcga	2220
cctatctcg cgatctgtct attcggtca tccatagttg cctgactccc cgctgttag	2280
ataactacgta tacgggagggtt cttaccatct ggccccagtg ctgcaatgat accgcgagac	2340
ccacgctcac cggctccaga tttatcagca ataaaccacg cagccggaaag ggccgagcgc	2400
agaagtggtc ctgcaacttt atccgcctcc atccagttca ttaattttgtt ccggaaagct	2460
agagtaagta gttcgccagt taatagtttgc gcaacgttgc ttgcattgc tacaggatc	2520
gtgggtgtac gtcgtcggtt tggatggct tcattcagct cgggtccca acgtcaagg	2580
cgagttacat gatccccat gttgtcaaa aaagcggtta gtccttcgg tcctccgatc	2640
gttgcgtacaa gtaagttggc cgcgttgc tcaactcatgg ttatggcagc actgcataat	2700
tctcttactg tcatgccatc cgtaagatgc tttctgtga ctggtagtgc ctcaaccaag	2760

US 10,337,001 B2

171

172

-continued

tcatttctgat	aatagtgtat	gcggcgaccg	agttgcctt	ccccggcgtc	aatacgggat	2820
aataccgcgc	cacatagc					2838
<210> SEQ ID NO 23						
<211> LENGTH: 2838						
<212> TYPE: DNA						
<213> ORGANISM: Artificial Sequence						
<220> FEATURE:						
<223> OTHER INFORMATION: synthetic construct						
<400> SEQUENCE: 23						
gcgtttctgg	gtgagcaaaa	acaggaaggc	aaaatgccgc	aaaaaaggga	ataagggcga	60
cacggaaatg	ttgaataactc	atactcttcc	ttttcaata	ttattgaagc	atttatcagg	120
gttattgtct	catgagcgga	tacatatttgc	aatgtatTTA	gaaaaataaa	caaataggGGG	180
ttccgcgcac	atTTTCCCGA	aaagtgcac	ctaaattgtA	agcgttaata	tttttgtaaa	240
attcgcgtta	aatTTTTGTT	aaatcagctc	atTTTTAAC	caataggccg	aaatcgccaa	300
aatcccTTAT	aaatcaaaaAG	aatagaccgA	gatagggttg	agtgtgttc	cagttggaa	360
caagAGTCCA	ctattaaAGA	acgtggactc	caacgtcaa	gggcgaaaaaa	ccgtctatca	420
gggcgatggc	ccactacgtg	aaccatcacc	ctaatacgt	ttttgggggt	cgagggtgccg	480
taaAGCACTA	aatcgGAACC	ctaaagggg	ccccggattt	agagcttgc	ggggaaAGCC	540
ggcgaacgtg	gcgagaaaagg	aagggaagaa	agcgaaagga	gccccgccta	ggcgctggc	600
aagtgttagcG	gtcacGCTGC	gcgtAAccAC	cacacccggc	gegttaatG	cgccgctaca	660
gggcgcgtcc	catTCGCCAT	tcaggctgCG	caactgttgg	gaaggggcgt	cggtgcgggc	720
cttttcgatTA	ttacgcEAGC	tggcgaaagg	gggatgtgot	gcaaggcgt	taagtgggt	780
aacGCCAGGG	ttttcccAGT	cacgacgttgc	taaaacgacg	gcaagtgc	gcgcgtata	840
cgactcacta	tagggcgaat	tgggtacgat	cgatgcggc	tggAACATT	ggtaattaaa	900
cttaacgcct	cagatttaga	cgaaggatttgc	aatggggaca	ttgtttatttgc	attctcgaa	960
gatacgcgcg	tgatATgcAG	ctccagcttt	tgttcccttgc	agtggggttt	aattgcgcgc	1020
ttggcgtaat	catggtcata	gctgtttcttgc	gtgtgaaatttgc	gttatccgct	cacaattccaa	1080
cacaacatac	gagccggaaag	cataaaagtgt	aaagcctggg	gtgcctaata	agttagctaa	1140
ctcacattaa	ttgcgttgCG	ctcaCTGCC	gctttccagt	cgggaaacct	gtcgtgcagg	1200
ctgcattaaat	gaatcgGCCA	acgcgcgggg	agaggcgggt	tgcgtattgg	gcccgtttcc	1260
gcttcctcgc	tcactgactc	gctgcgc	gtcggttgc	tgccggcagc	ggtatcagct	1320
cactcaaagg	cggtataacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatgt	1380
ttagccaaAG	gccagcaaaa	ggccagggaaac	cgtaaaaagg	ccgcgttgct	ggcgttttcc	1440
cataggctcc	ccccccctGA	cgagcatcac	aaaaatcgac	gtcaagtca	gagggtggcga	1500
aacCCGACAG	gactataaAG	ataccaggcg	ttttccctg	gaagctccct	cgtgcgtct	1560
cctgttccga	ccctgcgcgt	taccggatac	ctgtccgcct	tttccttc	gggaagcgtg	1620
gegcTTTCTC	atagctcAcG	ctgttaggtat	ctcagttcg	tgttaggtcg	tcgcgtcca	1680
ctgggcgttg	tgcacgaacc	ccccgttca	cccgaccgt	gegccttac	cggtaaactat	1740
cgtcttgcgt	ccaaACCCG	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	1800
aggattagca	gagcgcggat	tgttaggcgt	gctacagagt	tcttgaagtg	gtggccta	1860
tacggctaca	ctagaaggac	agtattttgtt	atctgcgtc	tgtgtaaagcc	agttaccttc	1920
ggaaaaaAGAG	ttggtagctc	ttgtatccggc	aaacaAAACCA	ccgctggtag	cggtgggttt	1980

US 10,337,001 B2

173

174

-continued

tttgttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tccttgatc	2040
tttctacgg ggtctgacgc tcagtggAAC gaaaactcac gttaagggt tttggtcATG	2100
agattatcaa aaaggatctt cacctagatc ctttAAATT aaaaatgaag tttaaatca	2160
atctaaagta tatatgagta aacttggct GACAGTtacc AATGCTTAAT CAGTGAGGCA	2220
cctatctcg cgatctgtct atttcggtca tccatagttg cctgactccc cgtcgtag	2280
ataactacga tacgggaggg cttaccatct ggcccAGT CGCAATGAT ACCGCGAGAC	2340
ccacgctcac cggctccaga tttatcagca ataaaccAGC CAGCCGGAAAG GGCCGAGCG	2400
agaagtggtc ctgcaacttt atccgcctcc atccAGTcta ttaattgtg ccgggaaGCT	2460
agagtaagta gttcgccagt taatagtttG CGCAACGTTG TTGCCATTGc tacaggcatc	2520
gtgggtgtcac gctcgctgtt tggtatggct tcattcagct cccggttccca acgtcaagg	2580
cgagttacat gatccccat gttgtgcaaa aaagcggtaa gtccttcgg tcctccgatc	2640
gttgcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcAGC actgcataat	2700
tctcttactg tcatgcatac cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag	2760
tcattctgag aatagtgtat gccccgaccc AGTTGCTTtT GccccggcGTC aatacgggat	2820
aataccgcgc cacatagc	2838

<210> SEQ ID NO 24

<211> LENGTH: 2838

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 24

gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcga	60
cacggaaatg ttgaataactc atacttcc ttttcaata ttattgaAGC atttatcagg	120
gttattgtct catgagcggA tacatatttG aatgtatttA gaaaaataAA caaataggGG	180
ttccggegcac attccccGA aaagtggccac ctaaattgtA agcgttaata ttttgttAAA	240
attcgegtta aattttgtt aaatcagtc attttttaac caataggccG aaatcgGAA	300
aatcccttat aaatcaaaaAG aatagacca gatagggttG agtgttGTC cagtttggAA	360
caagagtcca ctattaaAGA acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca	420
gggcgatggc ccactacgtG aaccatcacc ctaatcaagt tttttgggt cgagggtggc	480
taaagcacta aatcggaacc ctaaaggGAAG cccccgattt agagcttGAC ggggaaAGCC	540
ggcgaacgtG gcgagaaaAGG aagggaaAGG AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC	600
aagtgttagcG gtcacgctGc gctgttaaccAC cacacccGc GCGCTTAATG CGCGCTACA	660
gggcgegtcc cattcgccat tcaggctGcG caactgttG gaagggcgat cggtgccggc	720
ctcttcgcta ttacgcccAGC tggcgaaAGG gggatgtGCT GCAAGGGCGAT taagtgggt	780
aacgcccAGG tttcccAGt cacgacgttG taaaacgacG GCGAGTgAGC GCGCGTAATA	840
cgactcacta tagggcgaat tgggtacGAT CGATGCGGC TCGGAACGCT GGTGATTCT	900
cccaatgcct cagatttGAGA CGAAGGCTG AATGGGGATA ttatttactc cttctccAGt	960
gatgtgcgcG tgatatgcAG ctccAGCTT tggcccttt AGTGGGGTT AATTGCGCGC	1020
ttggcgtaat catggcatac GCTGTTCTt GTGTGAAATT GTTATCCGCT CACAATTCCA	1080
cacaacatac gagccggaAG cataaAGTGT AAAGCCTGGG GTGCCTAATG AGTGAAGCTAA	1140

US 10,337,001 B2

175**176**

-continued

ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cggaaacacct gtgtgcacag	1200
ctgcattaat gaatcgcca acgegcgggg agaggcgggt tgcgtattgg gcgccttc	1260
gcttcctcgc tcactgactc gctgcgcctcg gtcgttcggc tgcggcggagc ggtatcagct	1320
cactcaaagg cggtataacg gttatccaca gaatcagggg ataacgcagg aaagaacatg	1380
ttagccaaaag gccagcaaaa ggccaggaaac cgtaaaaagg ccgcgttgc ggcgttttc	1440
cataggctcc gccccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga	1500
aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgcct	1560
cctgttccga ccctgcgcct taccggatac ctgtccgcct ttctccctc gggaaagcgtg	1620
gcgcgttctc atagtcacg ctgttaggtat ctcaagttcgg tgcgttgcgt tcgcctcaag	1680
ctggggctgtg tgcaacgacc ccccggttcag cccgaccgcgt gcgccttatac cggtaaactat	1740
cgtcttgagt ccaacccggta aagacacgcac ttatcgccac tggcagcagc cactggtaac	1800
aggattagca gagcggaggta tgcgttgcgt gctacagagt tcttgaagtg gtggcctaacc	1860
tacggctaca ctagaaggac agtattttgtt atctgcgcctc tgctgaagcc agttaccc	1920
ggaaaaagag ttggtagctc ttgtatccggc aaacaaaccca ccgctggtag cggtgggttt	1980
tttggggcata aacggcggat tacggcggcata aaaaaaggat ctcaagaaga tcctttgatc	2040
ttttctacgg ggtctgacgc tcagttggAAC gaaaactcac gttaagggtt tttgggtcatg	2100
agattatcaa aaaggatctt cacctagatc ctttaaatt aaaaatgaag tttaaatca	2160
atctaaagta tatatgagta aacttggtct gacagttacc aatgcataat cagtggcga	2220
cctatctcag cgatctgtct atttcgttca tccatagttt cctgactccc cgctgtgt	2280
ataactacga tacgggagggtt cttaccatct ggccccagtg ctgcaatgtat accgcgagac	2340
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaaag ggccgagcgc	2400
agaagtggtc ctgcaacttt atccgcctcc atccagtctt ttaattttttt ccggaaagct	2460
agagtaagta gttcgttgcgt taatagtttgc caacacgttgc ttgcatttc tacaggcattc	2520
gtgggtgtcac gtcgtcggtt tggatggct tcattcagct ccgggttccca acgtcaagg	2580
cgagttacat gatccccat gttgtgcataaa aaagggttta gtccttcgg tccctccatc	2640
gttgcgttgcataa gtaagtggc cgcgtgtta tcaactcatgg ttatggcagc actgcataat	2700
tctcttactg tcatggccatc cgtaagatgc ttttctgtga ctgggttagt ctcaaccaag	2760
tcattctgag aatagtgtat gcccggaccc agttgttgc ttttgcgttgc aatacggat	2820
aataccgcgc cacatagc	2838

<210> SEQ ID NO 25
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 25

aguccucauc ucccucaagc guuuuagac uaugcuguuu ugaauggucc caaaac 56

<210> SEQ ID NO 26
<211> LENGTH: 86
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 26

US 10,337,001 B2

177

-continued

ggaacctauuc aaaacagcau agcaaguuua aauaaggcua guccguuauc aacuuguaaa	60
aguggcaccg agucggugcu uuuuuu	86

<210> SEQ ID NO 27
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 27

ugguaaugau ggcuacaaca guuuuagagc uaugcuguuu ugaauggucc caaaac

56

<210> SEQ ID NO 28
<211> LENGTH: 86
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 28

ggaacctauuc aaaacagcau agcaaguuua aauaaggcua guccguuauc aacuuguaaa

60

aguggcaccg agucggugcu uuuuuu

86

<210> SEQ ID NO 29
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 29

aguuccuauc ucccuacaagc guuuuagagc uaugcuguuu ugaauggucc caaaac

56

<210> SEQ ID NO 30
<211> LENGTH: 86
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (57)..(57)
<223> OTHER INFORMATION: u is 5-(3-Aminoallyl)-uridine-5'-triphosphate,
labeled with Cyanine5

<400> SEQUENCE: 30

ggaacctauuc aaaacagcau agcaaguuua aauaaggcua guccguuauc aacuuguaaa

60

aguggcaccg agucggugcu uuuuuu

86

<210> SEQ ID NO 31
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-methyl
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-methyl
<222> LOCATION: (53)..(55)

<400> SEQUENCE: 31

ugguaaugau ggcuacaaca guuuuagagc uaugcuguuu ugaauggucc caaaac

56

178

US 10,337,001 B2

179**180**

-continued

```

<210> SEQ ID NO 32
<211> LENGTH: 86
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl
<222> LOCATION: (83)..(85)

<400> SEQUENCE: 32
ggAACCAUUC AAAACAGCAU AGCAAGUUUA AAUAAGGCUA GUCCGUUAUC AACUUGUAAA 60
aguggcaccg agucggugcu uuuuuu 86

```

```

<210> SEQ ID NO 33
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl
<222> LOCATION: (53)..(55)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (53)..(56)

<400> SEQUENCE: 33
ugguaaugau ggcuuaaca guuuuagac uaugcuguuu ugaauggucc caaaac 56

```

```

<210> SEQ ID NO 34
<211> LENGTH: 86
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (83)..(85)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (83)..(86)

<400> SEQUENCE: 34
ggAACCAUUC AAAACAGCAU AGCAAGUUUA AAUAAGGCUA GUCCGUUAUC AACUUGUAAA 60
aguggcaccg agucggugcu uuuuuu 86

```

```

<210> SEQ ID NO 35
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:

```

US 10,337,001 B2

181**182**

-continued

```

<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (53)..(55)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (53)..(56)

```

<400> SEQUENCE: 35

ugguaaugau ggcuucaaca guuuuagagc uaugcuguuu ugaauggucc caaaac

56

```

<210> SEQ ID NO 36
<211> LENGTH: 86
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (83)..(85)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (83)..(86)

```

<400> SEQUENCE: 36

ggaaccuuuc aaaacagcau agcaaguua aauaaggcua guccguuauc aacuuguaaa

60

aguggcaccg agucggugcu uuuuuu

86

```

<210> SEQ ID NO 37
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (55)..(55)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (55)..(56)

```

<400> SEQUENCE: 37

ugguaaugau ggcuucaaca guuuuagagc uaugcuguuu ugaauggucc caaaac

56

```

<210> SEQ ID NO 38
<211> LENGTH: 86
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)

```

US 10,337,001 B2

183**184**

-continued

```

<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (85)..(85)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (85)..(86)

<400> SEQUENCE: 38

ggAACCAUUC AAAACAGCAU AGCAAGUUUA AAUAAGGCUA GUCCGUUAUC AACUUGUAAA      60
AGGGGACCG AGUCGGUGCU UUUUUU                                         86

<210> SEQ ID NO 39
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2-thiouridine nucleotide
<222> LOCATION: (3)..(3)

<400> SEQUENCE: 39

AGGUCCUCAUC UCCCUCUAGC GUUUAAGAGC UAAUGCUGUUU UGAUAUGGUCC CAAAAC      56

<210> SEQ ID NO 40
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2-thiouridine nucleotide
<222> LOCATION: (9)..(9)

<400> SEQUENCE: 40

AGGUCCUCAUC UCCCUCUAGC GUUUAAGAGC UAAUGCUGUUU UGAUAUGGUCC CAAAAC      56

<210> SEQ ID NO 41
<211> LENGTH: 56
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2-thiouridine nucleotide
<222> LOCATION: (11)..(11)

<400> SEQUENCE: 41

AGGUCCUCAUC UCCCUCUAGC GUUUAAGAGC UAAUGCUGUUU UGAUAUGGUCC CAAAAC      56

<210> SEQ ID NO 42
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 42

AGGUCCUCAUC UCCCUCUAGC GUUUAAGAGC UAAUGCUGGUAA ACAGCAUAGC AAGUUUAAA      60
AAGGCUCAGUC CGUUUAUCAAC UUGAAAAAGU GGCACCGAGU CGGUGCUUUU UUU             113

<210> SEQ ID NO 43
<211> LENGTH: 100
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

```

US 10,337,001 B2

185

-continued

186

<400> SEQUENCE: 43
 aguccucauc ucccuucaagc guuuuagagc uaguaauagc aaguuaaaau aaggcuaguc 60
 cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> SEQ ID NO 44
 <211> LENGTH: 113
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 44
 gcagauguag uguuuuccaca guuuuagagc uaugcuggaa acagcauagc aaguuaaaau 60
 aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

<210> SEQ ID NO 45
 <211> LENGTH: 111
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 45
 uccucaucuc ccuucaagcgu uuaagagcua ugcugguaac agcauagcaa guuuuaauaa 60
 ggcuaguuccg uuaucacuu gaaaaagugg caccgagucg gugcuuuuuu u 111

<210> SEQ ID NO 46
 <211> LENGTH: 110
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 46
 ccucaucucc cucaaggcgu uaagagcuaa gcugguaaca gcauagcaag uuuaauaaag 60
 gcuaguuccg uuaucacuuug aaaaaguggc accgagucgg ugcuuuuuuu 110

<210> SEQ ID NO 47
 <211> LENGTH: 114
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 47
 gaguccucau cucccuucaag cguuuuagag cuaugcugga aacagcauag caaguuaaa 60
 uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag ucggugcuuu uuuu 114

<210> SEQ ID NO 48
 <211> LENGTH: 115
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 48
 ggaguccuca ucuuccucaa gcguuuuaga gcuaugcugg uaacagcaua gcaaguuaaa 60
 auaaggcuag uccguuaucu acuugaaaaa guggcaccga gucggugcuu uuuu 115

<210> SEQ ID NO 49

US 10,337,001 B2

187

-continued

```

<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 49

aguccucauc ucccuacaagc guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aauucuaguc cguuuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu                  113

<210> SEQ ID NO 50
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 50

aguccucauc ucccuacaagc guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaaacuaguc cguuuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu                  113

<210> SEQ ID NO 51
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 51

aguccucauc ucccuacaagc guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaaacuaguu uguuuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu                  113

<210> SEQ ID NO 52
<211> LENGTH: 111
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 52

ggacuuuuuu uaguccuau cuccuacaag cguuuuagag cuagaaaaag caaguuaaaa      60
uaaggcuagu cgguuaucaa cuugaaaaag uggcaccgag ucggugcuuu u                  111

<210> SEQ ID NO 53
<211> LENGTH: 116
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 53

gaugaggacu uuuuuuaguc cuaucuccc ucaagcguuu uagagcuaga aaauagcaagu      60
aaaaauaagg cuaguccguu aucaacuuga aaaaguggca ccgagucggu gcuuuu                 116

<210> SEQ ID NO 54
<211> LENGTH: 111
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 54

guuuguuuuu uaguccuau cuccuacaag cguuuuagag cuagaaaaag caaguuaaaa      60

```

US 10,337,001 B2

189**190**

-continued

uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag ucggugcuuu u	111
--	-----

```

<210> SEQ ID NO 55
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: a is 5'-dimethoxytrityl-adenosine
<222> LOCATION: (1)..(1)

```

<400> SEQUENCE: 55

aguuccuauc ucccucaagc guuuuagagc uaugcuggua acagcauagc aaguuuuaau	60
---	----

aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 56
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

```

<400> SEQUENCE: 56

aguuccuauc ucccucaagc guuuuagagc uaugcuggua acagcauagc aaguuuuaau	60
---	----

aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 57
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (39)..(39)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (39)..(39)
<223> OTHER INFORMATION: n is a, c, g, or u, unknown or other

```

<400> SEQUENCE: 57

aguuccuauc ucccucaagc guuuuagagc uaugcuggna acagcauagc aaguuuuaau	60
---	----

aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 58
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (39)..(39)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (39)..(39)
<223> OTHER INFORMATION: n is a, c, g, or u, unknown or other
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl
<222> LOCATION: (110)..(112)

```

<400> SEQUENCE: 58

aguuccuauc ucccucaagc guuuuagagc uaugcuggna acagcauagc aaguuuuaau	60
---	----

US 10,337,001 B2

191

192

-continued

aaggcuauaguc cguaauacaac uugaaaaagu ggcaccgagu cggugcuiuu uuu 113

```
<210> SEQ ID NO 59
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (39)..(39)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (39)..(39)
<223> OTHER INFORMATION: n is a, c, g, or u, unknown or other
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (110)..(112)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (110)..(113)
```

<400> SEQUENCE: 59

aguccucauc ucccucaaggc guuuuaagagc uauggcgggna acagcauaggc aaguuuuauau

aaggcuauaguc cguuaaucaac uugaaaaagu ggcaccggagu cggugcuuuu uuu 113

```
<210> SEQ ID NO 60
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (39)..(39)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (39)..(39)
<223> OTHER INFORMATION: n is a, c, g, or u, unknown or other
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (109)..(111)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (109)..(112)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (110)..(112)
```

<400> SEQUENCE: 60

агууссугаагс үггүүсүүзээдс сүүшүүзээдс үнүүсүүзээрээ үгээсгүүзээдс үзүүшүүзээдс

<210> SEQ ID NO: 61
<211> LENGTH: 102
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

US 10,337,001 B2

193

194

-continued

```

<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (2)..(2)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2)..(2)
<223> OTHER INFORMATION: n is a, c, g, or u, unknown or other
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (68)..(68)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (68)..(68)
<223> OTHER INFORMATION: n is a, c, g, or u
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (100)..(100)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (100)..(100)
<223> OTHER INFORMATION: n is a, c, g, or u
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (101)..(101)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (101)..(102)

```

<400> SEQUENCE: 61

ungcagaugu aguguuucca caguuuaaga gcuaguauua gcaaguuuaa auaggcuag	60
uccguuanca acuugaaaaa guggcaccga gucgugucun uu	102

```

<210> SEQ ID NO 62
<211> LENGTH: 100
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (34)..(34)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (34)..(34)
<223> OTHER INFORMATION: n is a, c, g, or u, unknown or other
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (74)..(74)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (74)..(74)
<223> OTHER INFORMATION: n is a, c, g, or u, unknown or other
<220> FEATURE:
<221> NAME/KEY: fluorophore
<222> LOCATION: (90)..(90)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (90)..(90)
<223> OTHER INFORMATION: n is a, c, g, or u, unknown or other
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (99)..(99)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage

```

US 10,337,001 B2

195**196**

-continued

<222> LOCATION: (99)..(100)

<400> SEQUENCE: 62

gcagauuguag uguuuuccaca guuuuaagagc uagnaaauagc aaguuuuaau aaggcuaguc	60
cguuaaucaac uugnaaaagu ggcaccgagn cggugcuuuu	100

<210> SEQ ID NO 63
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(3)

<400> SEQUENCE: 63

aguuccucauc ucccucaagc guuuuaagagc uaugcuggaa acagcauagc aaguuuuaau	60
aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113

<210> SEQ ID NO 64
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(4)

<400> SEQUENCE: 64

aguuccucauc ucccucaagc guuuuaagagc uaugcuggaa acagcauagc aaguuuuaau	60
aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113

<210> SEQ ID NO 65
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(5)

<400> SEQUENCE: 65

aguuccucauc ucccucaagc guuuuaagagc uaugcuggaa acagcauagc aaguuuuaau	60
aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113

<210> SEQ ID NO 66
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (109)..(113)

<400> SEQUENCE: 66

aguuccucauc ucccucaagc guuuuaagagc uaugcuggaa acagcauagc aaguuuuaau	60
aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113

<210> SEQ ID NO 67
<211> LENGTH: 113

US 10,337,001 B2

197**198**

-continued

```

<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (20)..(20)

<400> SEQUENCE: 67
aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau 60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

<210> SEQ ID NO 68
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (19)..(19)

<400> SEQUENCE: 68
aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau 60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

<210> SEQ ID NO 69
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (18)..(18)

<400> SEQUENCE: 69
aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau 60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

<210> SEQ ID NO 70
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (17)..(17)

<400> SEQUENCE: 70
aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau 60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

<210> SEQ ID NO 71
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (17)..(18)

<400> SEQUENCE: 71
aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau 60

```

US 10,337,001 B2

199**200**

-continued

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 72
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (110)..(112)

<400> SEQUENCE: 72

```

aguccuaucc ucccuucaagc guuuuagc uaugcuggua acagcauagc aaguuuuaau	60
--	----

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 73
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (110)..(112)

<400> SEQUENCE: 73

```

gcagagaug uguuuccaca guuuuagc uaugcuggaa acagcauagc aaguuuuaau	60
--	----

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 74
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(20)

<400> SEQUENCE: 74

```

aguccuaucc ucccuucaagc guuuuagc uaugcuggua acagcauagc aaguuuuaau	60
--	----

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 75
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(26)

<400> SEQUENCE: 75

```

aguccuaucc ucccuucaagc guuuuagc uaugcuggua acagcauagc aaguuuuaau	60
--	----

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

<210> SEQ ID NO 76	
--------------------	--

US 10,337,001 B2

201**202**

-continued

```

<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) ..(37)

<400> SEQUENCE: 76

```

```

aguuccuaucc ucccuuagc guuuuagac uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu      113

```

```

<210> SEQ ID NO 77
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) ..(7)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (9) ..(11)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (13) ..(14)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (20) ..(20)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (30) ..(31)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (33) ..(33)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (35) ..(36)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (42) ..(42)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (45) ..(45)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (76) ..(77)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (80) ..(82)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (90) ..(90)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (95) ..(96)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (100) ..(101)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (104) ..(104)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (106) ..(112)

<400> SEQUENCE: 77

```

```

aguuccuaucc ucccuuagc guuuuagac uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu      113

```

```
<210> SEQ ID NO 78
```

US 10,337,001 B2

203

204

-continued

```

<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) ..(7)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (9) ..(11)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (17) ..(17)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (27) ..(28)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (30) ..(30)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (32) ..(33)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (36) ..(36)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (39) ..(39)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (42) ..(42)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (44) ..(44)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (47) ..(47)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (57) ..(57)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (62) ..(63)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (67) ..(68)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (73) ..(74)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (77) ..(78)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (87) ..(87)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (90) ..(90)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (92) ..(93)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (97) ..(98)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (101) ..(101)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (103) ..(109)

```

<400> SEQUENCE: 78

aguccucauc ucccucaagc guuuuagac uaugcuggua acagcauagc aaguuuuauu 60

aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

US 10,337,001 B2

205**206**

-continued

<210> SEQ_ID NO 79
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(7)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (9)..(11)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (17)..(17)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (27)..(28)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (30)..(30)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (32)..(33)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (36)..(36)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (39)..(39)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (42)..(42)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (44)..(44)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (47)..(47)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (57)..(57)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (62)..(63)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (67)..(68)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (73)..(74)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (77)..(78)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (87)..(87)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (90)..(90)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (92)..(93)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (97)..(98)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (101)..(101)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (103)..(109)

<400> SEQUENCE: 79

aguccucauc ucccucaaggc guuuuaagacg uaugcuggua acagcauagc aaguuuuaau

60

aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu

113

US 10,337,001 B2

207

208

-continued

```

<210> SEQ ID NO 80
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (8)..(8)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (17)..(18)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (21)..(21)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (26)..(26)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (28)..(29)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (32)..(32)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (34)..(34)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (37)..(38)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (40)..(41)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (43)..(44)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (46)..(46)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (48)..(49)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (51)..(52)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (58)..(58)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (61)..(64)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (67)..(67)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (78)..(79)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (83)..(89)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (91)..(92)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (94)..(94)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (97)..(99)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (102)..(103)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (105)..(105)

```

US 10,337,001 B2

209**210**

-continued

```
<400> SEQUENCE: 80
aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau 60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113
```

```
<210> SEQ_ID NO 81
<211> LENGTH: 100
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (8)..(8)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (17)..(18)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (21)..(21)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (26)..(26)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (28)..(29)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (32)..(33)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (35)..(36)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (38)..(39)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (41)..(42)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (46)..(46)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (48)..(48)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (51)..(54)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (57)..(57)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (68)..(69)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (73)..(79)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (81)..(82)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (84)..(84)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (87)..(89)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (92)..(93)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (95)..(95)
```

<400> SEQUENCE: 81

US 10,337,001 B2

211**212**

-continued

aguccucauc ucccuacaagc guuuuagagc uaguaauagc aaguuaaaau aaggcuaguc	60
cguuaauaac uugaaaaaagu ggcaccgagu cggugcuuuu	100

```

<210> SEQ ID NO 82
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (110)..(112)

<400> SEQUENCE: 82

```

gcagauguag uguuuccaca guuuaagagc uaugcuggaa acagcauagc aaguuuaau	60
aaggcuaguc cguuaauaac uugaaaaaagu ggcaccgagu cggugcuuuu uuu	113

```

<210> SEQ ID NO 83
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (6)..(6)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (8)..(8)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (11)..(11)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (13)..(17)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (30)..(31)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (33)..(33)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (35)..(36)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (39)..(39)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (42)..(42)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (45)..(45)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (47)..(47)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (50)..(50)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (56)..(56)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (60)..(60)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide

```

US 10,337,001 B2

213

214

-continued

```

<222> LOCATION: (65)..(66)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (69)..(71)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (76)..(77)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (80)..(82)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (90)..(90)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (93)..(93)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (95)..(96)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (100)..(101)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (104)..(104)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (106)..(112)

<400> SEQUENCE: 83

gcagauguag uguuuuccaca guuuuagagc uaugcuggua acagcauagc aaguuuuauu      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu                  113

```

```

<210> SEQ ID NO 84
<211> LENGTH: 100
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(5)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (26)..(29)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (32)..(33)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (35)..(36)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (38)..(39)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (41)..(42)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (46)..(46)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (48)..(49)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (52)..(54)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (57)..(57)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (68)..(69)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (73)..(79)
<220> FEATURE:

```

US 10,337,001 B2

215**216**

-continued

```

<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (81)..(82)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (84)..(84)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (87)..(89)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (92)..(93)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (95)..(95)

<400> SEQUENCE: 84

gcagauguag uguuuuccaca guuuuagagc uaguaauagc aaguuaaaau aaggcuaguc      60
cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu                           100

<210> SEQ ID NO 85
<211> LENGTH: 100
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (6)..(6)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (8)..(8)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (11)..(11)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (13)..(17)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (25)..(25)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (30)..(31)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (34)..(34)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (37)..(37)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (40)..(40)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (50)..(50)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (55)..(56)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (59)..(61)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (66)..(67)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (70)..(72)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (80)..(80)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (83)..(83)

```

US 10,337,001 B2

217**218**

-continued

```

<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (85)..(86)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (90)..(91)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (94)..(94)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (96)..(99)

<400> SEQUENCE: 85
gcagaaugau gguuuuccaca guuuuagagc uaguaauagc aaguuaaaaau aaggcuaguc      60
cguaaucaac uugaaaaagu ggcaccgagu cggugcuuuu                               100

<210> SEQ ID NO 86
<211> LENGTH: 113
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-deoxy-nucleotide
<222> LOCATION: (1)..(20)

<400> SEQUENCE: 86
agtccatc tccctcaagc guuuuagagc uaugcuggua acagcauagc aaguuaaaaau      60
aaggcuaguc cguaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuuu             113

<210> SEQ ID NO 87
<211> LENGTH: 113
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-deoxy-nucleotide
<222> LOCATION: (1)..(26)

<400> SEQUENCE: 87
agtccatc tccctcaagc gtttaagagc uaugcuggua acagcauagc aaguuaaaaau      60
aaggcuaguc cguaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuuu             113

<210> SEQ ID NO 88
<211> LENGTH: 113
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-deoxy-nucleotide
<222> LOCATION: (1)..(37)

<400> SEQUENCE: 88
agtccatc tccctcaagc gtttaagagc tatgctggua acagcauagc aaguuaaaaau      60
aaggcuaguc cguaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuuu             113

<210> SEQ ID NO 89
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-deoxy-nucleotide

```

US 10,337,001 B2

219**220**

-continued

```

<222> LOCATION: (15)..(15)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (15)..(16)

<400> SEQUENCE: 89

gcagauguag uguuuuccaca guuuuaagac uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuumu uuu                  113

<210> SEQ ID NO 90
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)

<400> SEQUENCE: 90

aguuccuaucc ucccuuacgc guuuuaagac uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuumu uuu                  113

<210> SEQ ID NO 91
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(3)

<400> SEQUENCE: 91

aguuccuaucc ucccuuacgc guuuuaagac uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuumu uuu                  113

<210> SEQ ID NO 92
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(4)

<400> SEQUENCE: 92

aguuccuaucc ucccuuacgc guuuuaagac uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuumu uuu                  113

<210> SEQ ID NO 93
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:

```

US 10,337,001 B2

221**222**

-continued

```

<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(5)

<400> SEQUENCE: 93

aguccucauc ucccucaagg guuuuagac uaugcuggua acagcauagc aaguuuuaau 60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

```

```

<210> SEQ ID NO 94
<211> LENGTH: 115
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(5)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(6)

<400> SEQUENCE: 94

ggaguccuca ucucccucaa gcguuuuaga gcuaugcugg uaacagcaua gcaaguuua 60
auaaggcuag uccguuauc acuugaaaa guggcaccga gucggugcuu uuuuu 115

```

```

<210> SEQ ID NO 95
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (109)..(112)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (109)..(113)

<400> SEQUENCE: 95

aguccucauc ucccucaagg guuuuagac uaugcuggua acagcauagc aaguuuuaau 60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

```

```

<210> SEQ ID NO 96
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (108)..(112)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (108)..(113)

<400> SEQUENCE: 96

```

```

aguccucauc ucccucaagg guuuuagac uaugcuggua acagcauagc aaguuuuaau 60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

```

```

<210> SEQ ID NO 97
<211> LENGTH: 114
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

```

US 10,337,001 B2

223**224**

-continued

```

<220> FEATURE:
<221> NAME/KEY: 5'-overhang (5' to the 20-nt guide sequence)
<222> LOCATION: (1) .. (1)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (3)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (4)

<400> SEQUENCE: 97

caguccucau cucccuacaag cguuuuaagag cuaugcuggu aacagcauag caaguuuaaa      60
uaaggcuagu ccguuaucaa cuugaaaaag uggcacccgag ucggugcuum uuuu                  114

<210> SEQ ID NO 98
<211> LENGTH: 114
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 5'-overhang (5' to the 20-nt guide sequence)
<222> LOCATION: (1) .. (1)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (3)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (4)

<400> SEQUENCE: 98

gaguccucau cucccuacaag cguuuuaagag cuaugcuggu aacagcauag caaguuuaaa      60
uaaggcuagu ccguuaucaa cuugaaaaag uggcacccgag ucggugcuum uuuu                  114

<210> SEQ ID NO 99
<211> LENGTH: 115
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 5'-overhang (5' to the 20-nt guide sequence)
<222> LOCATION: (1) .. (2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (5)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (6)

<400> SEQUENCE: 99

ucaguccuca ucuccccucaa gcguuuuaaga gcuaugcugg uaacagcaua gcaaguuuaa      60
uaaggcuagu uccguuaucaa acuugaaaaa guggcacccgag ucggugcuum uuuu                  115

<210> SEQ ID NO 100
<211> LENGTH: 115
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 5'-overhang (5' to the 20-nt guide sequence)
<222> LOCATION: (1) .. (2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (5)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (6)

```

US 10,337,001 B2

225**226**

-continued

<400> SEQUENCE: 100
 agaguccuca ucucccucaa gcguuuuaaga gcuaugcugg uaacagcaua gcaaguuuuaa 60
 auaaggcuag uccguuaauca acuugaaaaa guggcaccga gucgugcuu uuuuu 115

<210> SEQ ID NO 101
 <211> LENGTH: 116
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic construct
 <220> FEATURE:
 <221> NAME/KEY: 5'-overhang (5' to the 20-nt guide sequence)
 <222> LOCATION: (1)..(3)
 <220> FEATURE:
 <221> NAME/KEY: 2'-O-Methyl nucleotide
 <222> LOCATION: (1)..(7)
 <220> FEATURE:
 <221> NAME/KEY: phosphonoacetate internucleotide linkage
 <222> LOCATION: (1)..(8)
 <220> FEATURE:
 <221> NAME/KEY: 2'-O-Methyl nucleotide
 <222> LOCATION: (112)..(115)
 <220> FEATURE:
 <221> NAME/KEY: phosphonoacetate internucleotide linkage
 <222> LOCATION: (112)..(116)

<400> SEQUENCE: 101
 cucaguccuc aucucccuca agcguuuuaag agcuaugcug guaacagcau agcaaguuuua 60
 aaaaaggcua guccguuauc aacuugaaaa aguggcaccg agucggugcu uuuuuu 116

<210> SEQ ID NO 102
 <211> LENGTH: 113
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic construct
 <220> FEATURE:
 <221> NAME/KEY: 2'-O-Methyl nucleotide
 <222> LOCATION: (20)..(20)
 <220> FEATURE:
 <221> NAME/KEY: phosphonoacetate internucleotide linkage
 <222> LOCATION: (20)..(21)

<400> SEQUENCE: 102
 aguccucauc ucccuaaagc guuuuagagc uaugcuggua acagcauagc aaguuuuau 60
 aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

<210> SEQ ID NO 103
 <211> LENGTH: 113
 <212> TYPE: RNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic construct
 <220> FEATURE:
 <221> NAME/KEY: 2'-O-Methyl nucleotide
 <222> LOCATION: (19)..(20)
 <220> FEATURE:
 <221> NAME/KEY: phosphonoacetate internucleotide linkage
 <222> LOCATION: (19)..(20)

<400> SEQUENCE: 103
 aguccucauc ucccuaaagc guuuuagagc uaugcuggua acagcauagc aaguuuuau 60
 aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu 113

<210> SEQ ID NO 104
 <211> LENGTH: 113
 <212> TYPE: RNA

US 10,337,001 B2

227**228**

-continued

```

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (18)..(18)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (18)..(19)

<400> SEQUENCE: 104

aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu                  113

<210> SEQ ID NO 105
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (17)..(17)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (17)..(18)

<400> SEQUENCE: 105

aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu                  113

<210> SEQ ID NO 106
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (17)..(18)
<220> FEATURE:
<221> NAME/KEY: phosphonoacetate internucleotide linkage
<222> LOCATION: (17)..(19)

<400> SEQUENCE: 106

aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu                  113

<210> SEQ ID NO 107
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (110)..(112)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (110)..(113)

<400> SEQUENCE: 107

aguccucauc ucccucaaggc guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60

```

US 10,337,001 B2

229**230**

-continued

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 108
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (108)..(112)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (108)..(113)

```

<400> SEQUENCE: 108	60
---------------------	----

gcagauguag uguuuccaca guuuuagagc uaugcuggaa acagcauagc aaguuuuaau	60
---	----

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 109
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(5)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (1)..(6)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (108)..(112)
<220> FEATURE:
<221> NAME/KEY: phosphorothioate internucleotide linkage
<222> LOCATION: (108)..(113)

```

<400> SEQUENCE: 109	60
---------------------	----

gcagauguag uguuuccaca guuuuagagc uaugcuggaa acagcauagc aaguuuuaau	60
---	----

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 110
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (110)..(112)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (110)..(113)

```

<400> SEQUENCE: 110	60
---------------------	----

aguccucauc ucccucaagc guuuuagagc uaugcuggua acagcauagc aaguuuuaau	60
---	----

US 10,337,001 B2

231**232**

-continued

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 111
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (112) .. (112)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (112) .. (113)

```

<400> SEQUENCE: 111

aguccuauc ucccuacaagc guuuuagc uaugcuggua acagcauagc aaguuuuaau	60
---	----

aaggcuaguc cguuaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113
---	-----

```

<210> SEQ ID NO 112
<211> LENGTH: 75
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (3)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (4)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (72) .. (74)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (72) .. (75)

```

<400> SEQUENCE: 112

aguccuauc ucccuacaagc guuuuagc uaugcuggua acagcauagc aaguuuuaau	60
---	----

aaggcuaguc cguuu	75
------------------	----

```

<210> SEQ ID NO 113
<211> LENGTH: 77
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (76) .. (76)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (76) .. (77)

```

<400> SEQUENCE: 113

aguccuauc ucccuacaagc guuuuagc uaugcuggua acagcauagc aaguuuuaau	60
---	----

aaggcuaguc cguuauc	77
--------------------	----

US 10,337,001 B2

233

234

-continued

```

<210> SEQ ID NO 114
<211> LENGTH: 78
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (77)..(77)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (77)..(78)

```

<400> SEQUENCE: 114

gaguccucau cucccuacaag cguuuuaagag cuaugcuggu aacagcauag caaguuuaaa	60
uaaggcuagu ccguuauc	78

```

<210> SEQ ID NO 115
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(3)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(4)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (110)..(112)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (110)..(113)

```

<400> SEQUENCE: 115

gcagauguag uguuuuccaca guuuuaagac uaugcuggaa acagcauagc aaguuuaaaau	60
aaggcuaguc cgguaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113

```

<210> SEQ ID NO 116
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (112)..(112)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (112)..(113)

```

<400> SEQUENCE: 116

gcagauguag uguuuuccaca guuuuaagac uaugcuggaa acagcauagc aaguuuaaaau	60
aaggcuaguc cgguaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	113

US 10,337,001 B2

235**236**

-continued

```

<210> SEQ ID NO 117
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 117
gauguugucg augaaaaagu guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuumu uuu                  113

<210> SEQ ID NO 118
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<221> NAME/KEY: 2-aminoadenosine
<222> LOCATION: (15)...(15)

<400> SEQUENCE: 118
gauguugucg augaaaaagu guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuumu uuu                  113

<210> SEQ ID NO 119
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

<400> SEQUENCE: 119
gauuuagacg aaggauugaa guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuumu uuu                  113

<210> SEQ ID NO 120
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2-aminoadenosine
<222> LOCATION: (15)...(15)

<400> SEQUENCE: 120
gauuuagacg aaggauugaa guuuuagagc uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuumu uuu                  113

<210> SEQ ID NO 121
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (6)...(6)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (8)...(8)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (11)...(11)

```

US 10,337,001 B2

237

238

-continued

```

<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (13)..(15)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (31)..(31)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (33)..(33)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (36)..(36)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (39)..(39)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (47)..(47)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (81)..(82)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (90)..(90)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (104)..(104)
<220> FEATURE:
<221> NAME/KEY: 5-methylUridine
<222> LOCATION: (107)..(112)

<400> SEQUENCE: 121

```

```

gcagauguag uguuuuccaca guuuaagac uaugcuggua acagcauagc aaguuuuaau      60
aaggcuaguc cguuaucac uugaaaaagu ggcaccgagu cggugcuuuu uuu

```

113

```

<210> SEQ ID NO 122
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: Z base
<222> LOCATION: (70)..(71)
<223> OTHER INFORMATION: Z base
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (70)..(71)
<223> OTHER INFORMATION: n is a, c, g, or u, or unknown or other

<400> SEQUENCE: 122

```

```

aguuccuaucc cuccucaagc guuuaagac uaugcuggua acagcauagc aaguuuuaau      60
aaggcuagun nguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu

```

113

```

<210> SEQ ID NO 123
<211> LENGTH: 100
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl-nucleotide
<222> LOCATION: (54)..(54)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl-nucleotide
<222> LOCATION: (57)..(57)

```

US 10,337,001 B2

239**240**

-continued

<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (99)..(99)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (99)..(100)

<400> SEQUENCE: 123

aguuccuaucc ucccuuagc guuuuagac uaguauauc aaguuuaau aagguaauc	60
cguuaucac aagaaauugu ggcaccgagu cggugcuuuu	100

<210> SEQ ID NO 124
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1)..(1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1)..(2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl-nucleotide
<222> LOCATION: (64)..(64)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl-nucleotide
<222> LOCATION: (67)..(67)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (112)..(112)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (112)..(113)

<400> SEQUENCE: 124

aguuccuaucc ucccuuagc guuuuagac uaugcuggua acagcauagc aaguuuaau	60
aagguaauc cguuaucac aagaaauugu ggcaccgagu cggugcuuuu uuu	113

<210> SEQ ID NO 125
<211> LENGTH: 113
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: Z base
<222> LOCATION: (95)..(96)
<223> OTHER INFORMATION: Z base
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (95)..(96)
<223> OTHER INFORMATION: n is a, c, g, or u, or unknown or other

<400> SEQUENCE: 125

aguuccuaucc ucccuuagc guuuuagac uaugcuggua acagcauagc aaguuuaau	60
aaggcuaguc cguuaucac uugaaaaagu ggcannagau cggugcuuuu uuu	113

<210> SEQ ID NO 126
<211> LENGTH: 74
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

US 10,337,001 B2

241**242**

-continued

```

<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (73) .. (73)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (73) .. (74)

```

<400> SEQUENCE: 126

aguccucauc ucccuuagc guuuuagc uaugcuggua acagcauagc aaguuuuaau	60
aaggcuaguc cguu	74

```

<210> SEQ ID NO 127
<211> LENGTH: 75
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (1) .. (1)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (1) .. (2)
<220> FEATURE:
<221> NAME/KEY: 2'-O-Methyl nucleotide
<222> LOCATION: (74) .. (74)
<220> FEATURE:
<221> NAME/KEY: thiophosphonoacetate internucleotide linkage
<222> LOCATION: (74) .. (75)

```

<400> SEQUENCE: 127

aguccucauc ucccuuagc guuuuagc uaugcuggua acagcauagc aaguuuuaau	60
aaggcuaguc cguua	75

```

<210> SEQ ID NO 128
<211> LENGTH: 50
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

```

<400> SEQUENCE: 128

ttatatgaac ataactcaat ttgtaaaaaa gggtattggg gaattcatta	50
--	----

```

<210> SEQ ID NO 129
<211> LENGTH: 50
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

```

<400> SEQUENCE: 129

aataatacttg tattgagttt aacatTTTTT cccataacc cttaaagtaat	50
---	----

```

<210> SEQ ID NO 130
<211> LENGTH: 98
<212> TYPE: RNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic construct

```

US 10,337,001 B2

243

244

-continued

<400> SEQUENCE: 130

auaacucaau uuguaaaaaa guuuuagac uauagcaagu uaaaaauaagg uaguccguua	60
ucaacuugaa aaaguggcac cgagucggug cuuuuuuu	98

We claim:

1. A synthetic CRISPR guide RNA having at least one 5'-end and at least one 3'-end, the synthetic guide RNA comprising:

- (a) one or more modified nucleotides within five nucleotides from said 5'-end, or
- (b) one or more modified nucleotides within five nucleotides from said 3'-end, or
- (c) both (a) and (b);

wherein said guide RNA comprises one or more RNA molecules, and has gRNA functionality comprising associating with a Cas protein and targeting the gRNA:Cas protein complex to a target polynucleotide, wherein the modified nucleotide has a modification to a phosphodiester linkage, a sugar, or both.

2. The synthetic guide RNA of claim 1 wherein said guide RNA is a single guide RNA (sgRNA).

3. The synthetic guide RNA of claim 1 wherein said guide RNA comprises:

- (a) one modified nucleotide at the first nucleotide of said 5'-end, or
- (b) at least one modified nucleotide within two nucleotides from said 3'-end, or
- (c) both (a) and (b).

4. The synthetic guide RNA of claim 3 wherein said one modified nucleotide is independently selected from a 2'-O-methyl nucleotide, 2'-O-methyl-3'-phosphorothioate nucleotide, 2'-O-methyl-3'-phosphonoacetate nucleotide, and 2'-O-methyl-3'-phosphonothioacetate nucleotide.

5. The synthetic guide RNA of claim 1 wherein said modified nucleotide comprises a modified internucleotide linkage or a modified terminal phosphate group selected from an alkylphosphonate, a phosphonocarboxylate, a phosphonoacetate, a boranophosphonate, a phosphorothioate, a phosphonothioacetate, and a phosphorodithioate group.

6. The synthetic guide RNA of claim 1 wherein said guide RNA comprises:

- (a) one or more modified nucleotides within five nucleotides from said 5'-end, or
- (b) at least two consecutive modified nucleotides within five nucleotides from said 3'-end, or
- (c) both (a) and (b).

7. The synthetic guide RNA of claim 6 wherein said one or more modified nucleotides within five nucleotides from the 3'-end are each independently selected from a 2'-O-methyl nucleotide, 2'-O-methyl-3'-phosphorothioate nucleotide, 2'-O-methyl-3'-phosphonoacetate nucleotide, 2'-O-methyl-3'-phosphonothioacetate nucleotide, and a combination thereof.

8. The synthetic guide RNA of claim 1 wherein said guide RNA comprises a modified internucleotide linkage or a modified terminal phosphate group selected from a phosphonocarboxylate, a phosphonoacetate, and a phosphonothioacetate group.

9. The synthetic guide RNA of claim 1 wherein said one or more modified nucleotides comprise a 2'-O-methyl-3'-phosphorothioate nucleotide.

10. The synthetic guide RNA of claim 1 wherein said one or more modifications are independently selected from a 2'-deoxy, 2'-O-methyl, 3'-phosphorothioate, 3'-phosphonoacetate (PACE), 3'-phosphonothioacetate (thioPACE), and a combination thereof.

11. The synthetic guide RNA of claim 10 wherein said one or more, modified nucleotides comprise a 2'-O-methyl-3'-phosphonoacetate nucleotide or a 2'-O-methyl-3'-phosphonothioacetate nucleotide.

12. A synthetic CRISPR crRNA molecule comprising a 5' end, a 3'-end, and a guide sequence capable of hybridizing to a target polynucleotide, wherein said crRNA molecule comprises:

- (a) one or more modified nucleotides within five nucleotides from said 5'-end, or
- (b) one or more modified nucleotides within five nucleotides from said 3'-end, or
- (c) both (a) and (b),

wherein said synthetic crRNA molecule has gRNA functionality comprising associating with a Cas protein and targeting the gRNA:Cas protein complex to a target polynucleotide, and wherein the modified nucleotide has a modification to a phosphodiester linkage, a sugar, or both.

13. The crRNA molecule of claim 12 wherein said crRNA molecule comprises:

- (a) one modified nucleotide at the first nucleotide of said 5'-end, or
- (b) at least one modified nucleotide within two nucleotides from said 3'-end, or
- (c) both (a) and (b).

14. The crRNA molecule of claim 13 wherein said one modified nucleotide is independently selected from a 2'-O-methyl nucleotide, 2'-O-methyl-3'-phosphorothioate nucleotide, 2'-O-methyl-3'-phosphonoacetate nucleotide and 2'-O-methyl-3'-phosphonothioacetate nucleotide.

15. The crRNA molecule of claim 12 wherein said modified nucleotide comprises a modified internucleotide linkage or a modified terminal phosphate group selected from an alkylphosphonate, a phosphonocarboxylate, phosphonoacetate, a boranophosphonate, a phosphorothioate, a phosphonothioacetate, and a phosphorodithioate group.

16. The crRNA molecule of claim 12 wherein said crRNA comprises a modified internucleotide linkage or a modified terminal phosphate group selected from a phosphonocarboxylate, a phosphonoacetate, and a phosphonothioacetate group.

17. The crRNA molecule of claim 12, wherein said one or more modified nucleotides comprise a modification independently selected from a 2'-deoxy, 2'-O-methyl, 3'-phosphorothioate, 3'-phosphonoacetate (PACE), 3'-phosphonothioacetate (thioPACE), and a combination thereof.

18. The crRNA molecule of claim 12 wherein said one or more modified nucleotides comprise a 2'-O-methyl-3'-phosphorothioate nucleotide.

19. The crRNA molecule of claim 12 wherein said one or more modified nucleotides comprise a 2'-O-methyl-3'-

US 10,337,001 B2

245

phosphonoacetate nucleotide or a 2'-O-methyl-3'-phosphothioacetate nucleotide or a combination thereof.

20. The crRNA molecule of claim **12** wherein said modified nucleotides are each independently selected from a 2'-O-methyl nucleotide, 2'-O-methyl-3'-phosphorothioate nucleotide, 2'-O-methyl-3'-phosphonoacetate nucleotide and 2'-O-methyl-3'-phosphonothioacetate nucleotide.

21. A method for modifying a DNA sequence, regulating the expression of a gene of interest, or cleaving a target polynucleotide, the method comprising:

- selecting at least one target polynucleotide;
- providing at least one synthetic guide RNA of claim **1**;
- forming a gRNA:Cas protein complex comprising a Cas protein and the synthetic guide RNA; and
- contacting the target polynucleotide with the gRNA:Cas protein complex;

wherein said Cas protein is provided as a protein or is expressed from a polynucleotide encoding said Cas protein.

22. The method of claim **21**, wherein said guide RNA comprises:

- (a) one modified nucleotide at the first nucleotide of said 5'-end, or
- (b) at least one modified nucleotide within two nucleotides from said 3'-end, or
- (c) both (a) and (b).

23. The method of claim **22** wherein said modified nucleotides are each independently selected from a 2'-O-

246

methyl nucleotide, 2'-O-methyl-3'-phosphorothioate nucleotide, 2'-O-methyl-3'-phosphonoacetate nucleotide and 2'-O-methyl-3'-phosphonothioacetate nucleotide.

24. The method of claim **21** wherein said modified nucleotide comprises a modified internucleotide linkage or a modified terminal phosphate group selected from an alkyl-phosphonate, a phosphonocarboxylate, a phosphonoacetate, a boranophosphonate, a phosphorothioate, a phosphonothioacetate, and a phosphorodithioate group.

25. The method of claim **21** wherein said one or more modified nucleotides comprise a 2'-O-methyl-3'-phosphorothioate nucleotide.

26. The method of claim **21** wherein said one or more modified nucleotides comprise a 2'-O-methyl-3'-phosphonoacetate nucleotide or a 2'-O-methyl-3'-phosphonothioacetate nucleotide or a combination thereof.

27. The method of claim **21** wherein said forming of said gRNA:Cas protein complex is performed outside or inside a cell and said contacting is performed in a cell.

28. The method of claim **21** wherein said Cas protein is a Cas9 protein.

29. The method of claim **21** wherein said forming is performed outside a cell.

30. A set or a library comprising at least two guide RNAs of claim **1**.

* * * * *